

Universidad de Santiago de Compostela

Facultad de Medicina y Odontología

Departamento de Medicina

Hospital Clínico Universitario de Santiago de Compostela

Instituto de Investigación Sanitaria de Santiago

Laboratorio de Investigación 10

**FACTORES GENÉTICOS ASOCIADOS
CON PREDISPOSICIÓN A LUPUS
ERITEMATOSO SISTÉMICO:**

**Confirmación, efecto funcional, repercusión fenotípica y
relación con la susceptibilidad a la Artritis Reumatoide**

Tesis Doctoral

María de los Ángeles Suárez Gestal

2010

El Dr. Antonio González Martínez-Pedrayo, investigador principal del Laboratorio de Investigación 10 del Instituto de Investigación Sanitaria del Complejo Hospitalario Universitario de Santiago de Compostela, y el Dr. Juan Jesús Gómez-Reino Carnota, Jefe del Servicio de Reumatología del Hospital Clínico Universitario de Santiago de Compostela y Profesor Titular del Departamento de Medicina y Catedrático de la Universidad de Santiago de Compostela,

CERTIFICAN QUE:

El presente trabajo que lleva por título: *Factores genéticos asociados con predisposición a Lupus Eritematoso Sistémico: Confirmación, efecto funcional, repercusión fenotípica y relación con la susceptibilidad a la Artritis Reumatoide*, realizado por María de los Ángeles Suárez Gestal en el Departamento de Medicina de la Universidad de Santiago de Compostela bajo nuestra dirección, ha sido revisado y está en disposición de ser presentado para optar al grado de Doctora en Biología.

Fdo: Dr. Antonio González Martínez-Pedrayo Fdo: Dr. Juan J. Gómez-Reino Carnota

A mi hermana Merchi

Agradecimientos

El camino ha sido duro y largo, con unos momentos mejores que otros, pero todo este esfuerzo ha merecido la pena, no sólo por los objetivos alcanzados en mi trabajo, sino también porque durante estos años he conocido a gente maravillosa que, junto a mi familia y amigos me han apoyado desde el principio hasta el final de mi tesis.

En primer lugar, quiero darle las gracias a mi director de tesis, el Dr. Antonio González, por haberme dado la oportunidad de realizar este trabajo, por haberme animado y por tener tantas ideas y proyectos. Gracias también al Dr. Juan Gómez-Reino, mi codirector de tesis, por su apoyo y por permitirme realizar este trabajo en colaboración con el Servicio de Reumatología del Hospital Clínico Universitario de Santiago de Compostela.

También agradezco la cooperación de los pacientes de lupus y artritis reumatoide, de los controles y de los médicos e investigadores que nos han proporcionado las muestras de ADN, así como los que nos han ayudado en la puesta a punto de los estudios funcionales.

Quiero agradecerle a mis compañeros todo ese entusiasmo y buen rollo que transmiten, que han hecho que los días duros fuesen más llevaderos y que estos cinco años se me hayan pasado volando. A Isa por haberme enseñado tantas cosas, en lo profesional como en lo personal. Ha sido toda una mentora para mí y un modelo a seguir. A Julio, por su entusiasmo, su ayuda y su filosofía. A Servet, por su cariño desbordante. A Isabel, por su alegría y simpatía. A Pombo, porque con sus ocurrencias nos hemos echado unas risas en el laboratorio. A Rebe, por las conversaciones de camino a casa, su comprensión, su alegría y sus piropos, que se echan mucho de menos. A Manolo por toda su ayuda, que ha sido muchísima, y su disponibilidad. Por ser un genio aún no descubierto, por sus amplios conocimientos en toda una variedad de temas, por su orden caótico, y por haberme dejado escribirlo todo en mi libreta. A Elisa, por su alegría y optimismo, por sus *creazioni*, por las largas y buenas conversaciones que hemos mantenido y los viajes que hemos compartido. A Cristina, porque cada día me sorprende con una cosa nueva, por su elegancia y sus consejos, que me han ayudado tanto. A Chicha, por ser como es, sincera, divertida, despierta. Porque sus diálogos con Rudy me han hecho llorar de risa. A Rudy, porque con él es imposible ponerse serio. A Rocío por saber escuchar. A Samu, por la paciencia que tuvo cuando me enseñó algunas técnicas. A Bea (laboratorio 8), por su alegría y sus “premoniciones”, que ojalá

se cumplan. A Diego, porque con sus sustos he descargado mucha adrenalina. A Ana (laboratorio 7), por ser una estupenda compañera de viaje y por su risa contagiosa. A Sonia, por tantos trayectos juntos del trabajo a casa, siendo un gran apoyo en mis primeros años. A Ezequiel, por organizar las pachangas de futbito. A Turo, por su amistad y porque el sábado noche en Pontevedra le gusta tanto como a mí. A Carmen, por su ayuda, su disponibilidad y su buen hacer. A Pepe, porque ha sido un buen compañero. A Bea (laboratorio 9), Paula y Vero por su dulzura. Y muchas gracias a todos los demás: Miguel, Javi, Anna, Giuseppe y Morena (laboratorio 9), Ana y Patricia (laboratorio 8), Rober, Chus, Sandra, Vane, Pamela y Manuel (laboratorio 7), Marina, Patricia y María (laboratorio 11) y a otras personas con las que he coincidido en el laboratorio, como Marta, Yolanda, Inés y Aida (enfermera), así como a las nuevas caras, Aida, Ariadna, Ana y Mari, con las que espero tener tiempo de compartir buenos momentos. También quiero agradecerle a Carmen, Fran y Oreste su ayuda siempre que la he necesitado.

A mis compis de facultad, por su compañía, su compañerismo y porque nos lo hemos pasado muy bien comiendo juntos los días que teníamos prácticas. A Bruno, porque además de un buen compañero es un fiel amigo, siempre ha estado ahí para escucharme y darme su opinión y para sacarme de fiesta con sus amigos cuando más lo necesitaba. A Raquel, por ser tan entrañable. A Javi R, porque sus rizos nos vuelven locas y por lo bien que encaja y devuelve los piques. A Javi S, por tener la investigación en la sangre y por su amor por el mar. Siempre me lo imaginaré con un neopreno y debajo del agua. Y no me olvido de Toño, Gonzalo, Raúl y Antonio, con los que he compartido buenos momentos y espero seguir haciéndolo. También quiero agradecerle a Jesús su ayuda generosa con el *photoshop*, que me ha sacado de un apuro; y a Fátima y Andrea, su cariño. Y a mis compañeros de los cursos de doctorado, especialmente a Javi, Galo y Ana, les agradezco su amistad, su apoyo y la alegría que me han transmitido.

A mis amigos de Santiago. A mi mejor amiga, Cristina, porque la conozco desde los 4 años y jamás nos hemos enfadado, porque siempre ha tenido unas palabras de ánimo, las necesitara o no, por lo mucho que me valora y porque se preocupa mucho por mí. A Raquel, porque en cuanto la conocí se convirtió en una de mis mejores amigas, porque su memoria compensaba mis despistes en la facultad y porque me descubrió el vicio por los lacasitos. A Toni, porque fue un buen compañero de pupitre y se ha convertido en un buen amigo. A Iván, por los buenos momentos que hemos compartido.

A mis niñas de Pontevedra, que me acogieron desde el primer momento como una más, la culpa de lo cual la tiene mi prima Rut. A ella le doy las

gracias por ser mi confidente y por aguantarme en los buenos y malos momentos. A Ana, por su sonrisa y por descubrirme grupos de música indies. A Lourdes, por su dulzura y por haberme escuchado cuando necesitaba desahogarme. A Iria S, porque con ella es fácil pasárselo bien. A Laura, por su cariño y sus trucos para el estrés. A Bea, porque me encanta cuando le afecta la altitud. A Carmen, por su originalidad y curiosidad. A Nieves, por su personalidad alegre. A Iria P, por mostrar interés cuando empiezo a hablar de cosas raras de mi tesis. A Susana, por su tranquilidad. Y a Vicky, por su energía.

Y siguiendo con más gente de Pontevedra, le quiero dar las gracias a Rita, porque en poco tiempo se ha convertido en una amiga en la que se puede confiar. A Peon y Claudio, por el buen rato que hemos pasado en las pachangas, en el rafting, en la carrera de orientación o tomando unas cañas. Y a Toño, porque durante un tiempo ha sido un gran apoyo para mí, hablar con él me tranquilizaba y me daba fuerzas para seguir. Por el cariño que me transmitió, porque los buenos momentos fueron muchos y esos son los que quedan en el recuerdo y porque sigo disfrutando de su compañía.

Y, por último, quiero agradecerle a mi familia todo su apoyo y su cariño. Especialmente, a mis padres, porque siempre me han animado a hacer todo lo que me proponía. A Belén, porque además de ser mi hermana pequeña ha sido mi amiga y con ella he vivido muchas “aventuras”. A Cris, por toda la paciencia que ha tenido cuando sus dos hermanas pequeñas no la dejábamos estudiar, porque me encanta ir de compras con ella y por preocuparse por mí. A Manu, por haberme dejado jugar con él, a pesar de llevarse siempre las culpas de bombillas y jarrones rotos, y porque sus notas me motivaron a intentar ser tan buena estudiante como él. A Jose, por su cariño y su paciencia cuando me pongo enferma y lo incordio con preguntas y favores. A mis sobrinos, Marcela, Álex, Javi, Marta, Jaime y Cris por mimarme tanto, por querer jugar conmigo y por considerarme una “tía guay”. A mis cuñados, Gaspar, Carmen e Isa, y a mi cuasi-cuñado Eloy, por haber querido formar parte de esta gran familia. Y a mis tíos Pepe y Pili, por su compañía en los días de fiesta y por todo su afecto y ayuda.

A todos vosotros gracias.

Prefacio

La presente tesis está estructurada en los apartados de introducción, hipótesis, objetivos, procedimientos experimentales y resultados, discusión y conclusiones. La introducción está enfocada en varios aspectos del lupus eritematoso sistémico (LES), incluyendo los mecanismos que lo causan, sus características clínicas, su genética y, dentro de este apartado, los factores genéticos de susceptibilidad más destacados. Por último, se exponen las evidencias existentes de que enfermedades autoinmunes, como son el LES y la artritis reumatoide (AR), comparten una parte de su componente genético. Los procedimientos experimentales y resultados se presentan en formato de artículos: tres ya han sido publicados, así como una carta al editor, mientras que otros tres han sido enviados a revisión. La discusión está dividida en capítulos que se estructuran de acuerdo con los artículos, terminando con una discusión general. Por último, se exponen las conclusiones a las que se ha llegado con este trabajo. Como anexo, se incluye un artículo enviado a revisión, que estudia predictores genéticos de la respuesta al tratamiento en AR.

ÍNDICE

Índice

Abreviaturas	1
Introducción	7
1. Lupus Eritematoso Sistémico	9
1.1. Patogenia del LES	10
1.2. Clínica del LES	14
1.3. El LES como enfermedad compleja	16
2. Influencia del componente genético en la variabilidad clínica del LES	17
3. Estudio de la genética del LES	19
3.1. Estudios de ligamiento	19
3.2. Estudios de asociación	21
3.2.1. <i>PDCDI</i>	23
3.3. Estudios de asociación de genoma completo	27
3.3.1. GWAS en LES	30
3.4. Factores genéticos de LES confirmados	37
3.4.1. <i>HLA</i>	37
3.4.2. Factores del complemento	40
3.4.3. <i>FCGRs</i>	42
3.4.4. <i>TREX1</i>	43
3.4.5. <i>IRF5</i>	45
3.4.6. <i>PTPN22</i>	46
3.4.7. <i>STAT4</i>	47
3.4.8. <i>ITGAM</i>	49
3.4.9. <i>C8orf13-BLK</i>	50
3.4.10. <i>BANK1</i>	51

3.4.11. <i>Iq25.1</i> y <i>TNFSF4</i>	52
3.4.12. <i>KIAA1542-IRF7</i>	53
3.4.13. <i>PXK</i>	53
3.4.14. <i>MECP2-IRAK1</i>	54
3.4.15. <i>TNFAIP3</i> y región intergénica <i>6q23</i>	55
3.4.16. <i>TNIP1</i>	57
3.4.17. <i>PRDMI-ATG5</i>	58
3.4.18. <i>JAZF1</i>	58
3.4.19. <i>UHRF1BP1</i>	59
3.4.20. <i>IL10</i>	59
3.5. Factores genéticos con asociación al LES controvertida	60
3.5.1. <i>TYK2</i>	60
3.5.2. <i>LY9</i>	61
4. Determinación de los efectos funcionales de los <i>loci</i> de susceptibilidad	62
5. El problema de la heredabilidad no explicada	66
6. La hipótesis de los factores genéticos comunes a múltiples enfermedades autoinmunes a la luz de los GWAS	70
6.1. <i>TNFAIP3</i> y región intergénica <i>6q23</i>	72
Hipótesis	75
Hipótesis	77
Objetivos	79
Objetivos	81
Procedimiento experimental y resultados	83
Esquema general	85
1. Factores genéticos de LES	89
Publicación 1	91
Publicación 2	105

2. Estudio funcional de PD1.3	107
Publicación 3	109
3. Influencia del componente genético en la variabilidad clínica del LES	117
Publicación 4	119
Publicación 5	151
4. Componente genético común a LES y AR	167
Publicación 6	169
Publicación 7	177
Discusión	199
1. Discusión	201
1.1. Confirmación de factores de susceptibilidad al LES	203
1.2. Análisis funcionales del SNP asociado a LES, PD1.3	213
1.3. Influencia del componente genético en la variabilidad clínica del LES	217
1.4. Componente genético común a LES y AR	225
2. Sumario	231
Conclusiones	235
Conclusiones	237
Bibliografía	239
Bibliografía	241
Anexo	277
Publicación 8	279

Abreviaturas

3'UTR	Región 3' no traducida (del inglés, <i>3' untranslated region</i>)
ACPA	Anticuerpos frente a proteínas/péptidos citrulinados (del inglés, <i>anti-citrullinated protein/peptide antibodies</i>)
ACR	Colegio Americano de Reumatología (del inglés, <i>American College of Rheumatology</i>)
ADN	Ácido desoxirribonucleico
AFF3	Miembro 3 de la familia AF4/FMR2 (del inglés, <i>AF4/FMR2 family member 3</i>)
ANA	Anticuerpo antinuclear (del inglés, <i>antinuclear antibody</i>)
anti-RNP	Anticuerpo anti-ribonucleoproteína
anti-Sm	Anticuerpo anti-Smith
aPL	Anticuerpo antifosfolípido
AR	Artritis reumatoide
ARN	Ácido ribonucleico
BAFF/BLyS	Factor activador de células B (del inglés, <i>B-cell activating factor/B-lymphocyte stimulator</i>)
BANK1	Proteína estructural de células B con repeticiones de anquirina (del inglés, <i>B-cell scaffold protein with ankyrin repeats</i>)
BCR	Receptor de células B (del inglés, <i>B-cell receptor</i>)
BLK	Quinasa de linfocito B (del inglés, <i>B-lymphocyte kinase</i>)
C3bi	Producto de activación del complemento 3
CARD15/NOD2	<i>Caspase recruitment domain 15/nucleotide-binding oligomerization domain containing 2</i>
CBF	<i>Core binding factor</i>
ChIP	Inmunoprecipitación de cromatina (del inglés, <i>chromatin immunoprecipitation</i>)
CNV	Variación del número de copias (del inglés, <i>copy number variation</i>)
CPA	Célula presentadora de antígeno

CREB1	Proteína 1 de unión al elemento de respuesta al cAMP (del inglés, <i>cAMP responsive element binding protein 1</i>)
CSK	Tirosín-quinasa Src C-terminal (del inglés, <i>C-terminal Src tyrosine kinase</i>)
CTLA4	Antígeno 4 de linfocitos T citotóxicos (del inglés, <i>cytotoxic T lymphocyte antigen 4</i>)
CV	Coefficiente de variación
DC	Célula dendrítica (del inglés, <i>dendritic cell</i>)
EIRA	Investigación Epidemiológica de Artritis Reumatoide (del inglés, <i>Epidemiological Investigation of Rheumatoid Arthritis</i>)
EMSA	Ensayo de cambio de movilidad electroforética (del inglés, <i>electrophoretic mobility shift assay</i>)
FCGR	Receptor de baja afinidad de la fracción constante de la inmunoglobulina G (del inglés, <i>Fc fragment of IgG low affinity receptor</i>)
GM-CSF	Factor estimulador de colonias de granulocitos-macrófagos (del inglés, <i>granulocyte-macrophage colony-stimulating factor</i>)
GRZB	Granzima B
GWAS	Estudio de asociación de genoma completo (del inglés, <i>genome wide association study</i>)
HIC2-UBE2L3	<i>Hypermethylated in cancer 2-ubiquitin-conjugating enzyme E2L 3</i>
HLA	Antígeno leucocitario humano (del inglés, <i>human leukocyte antigen</i>)
ICAM1	Molécula 1 de adhesión intracelular (del inglés, <i>intracellular adhesion molecule 1</i>)
IFN	Interferón
IFNAR	Receptor de interferón α (del inglés, <i>IFNα receptor</i>)
Ig	Inmunoglobulina
IL	Interleucina
IL1R	Receptor de interleucina 1 (del inglés, <i>IL1 receptor</i>)
IL2R	Receptor de interleucina 2 (del inglés, <i>IL2 receptor</i>)
indel	Polimorfismo del tipo inserción/delección

IP3R	Receptor de inositol 1,4,5-trifosfato (del inglés, <i>inositol 1,4,5-triphosphate receptor</i>)
IRAK1	Quinasa 1 asociada a receptor de interleucina 1 (del inglés, <i>interleukin 1 receptor associated kinase 1</i>)
IRF	Factor regulador del interferón (del inglés, <i>interferon regulatory factor</i>)
ITG	Integrina
ITIM	<i>Immunoreceptor tyrosine-based inhibition motif</i>
ITPR2	Receptor 2 de inositol 1,4,5-trifosfato (del inglés, <i>inositol 1,4,5-triphosphate receptor 2</i>)
ITSM	<i>Immunoreceptor tyrosine-based switch motif</i>
JAZF1	<i>JAZF zinc finger 1</i>
KO	<i>Knockout</i>
KYF5A	Miembro 5A de la familia de las quinesinas (del inglés, <i>kinesin family member 5A</i>)
LD	Desequilibrio de ligamiento (del inglés, <i>linkage disequilibrium</i>)
LES	Lupus eritematoso sistémico
LYN	<i>Yamaguchi sarcoma viral oncogene homolog</i>
LYP	Fosfatasa de tirosina linfoide (del inglés, <i>lymphoid-specific tyrosine phosphatase</i>)
MECP2	Proteína 2 de unión a CpG metilado (del inglés, <i>methyl-CpG-binding protein 2</i>)
MHC	Complejo mayor de histocompatibilidad (del inglés, <i>major histocompatibility complex</i>)
MIP1 α	Proteína 1 α inflamatoria de macrófagos (del inglés, <i>Macrophage inflammatory protein 1α</i>)
MMEL1	<i>Membrane metallo-endopeptidase-like 1</i>
Mona/Gads	Adaptador monocítico (del inglés, <i>monocytic adaptor</i>)
NARAC	Consorcio Norteamericano de Artritis Reumatoide (del inglés, <i>North American Rheumatoid Arthritis Consortium</i>)
NF κ B	Factor nuclear κ B (del inglés, <i>nuclear factor κB</i>)

NK	Células “asesinas naturales” (del inglés, <i>natural killer</i>)
NOD	Ratones diabéticos no obesos (del inglés, <i>non-obese diabetic</i>)
nsSNP	SNP no sinónimo
OLIG3	Factor de transcripción 3 de oligodendrocitos (del inglés, <i>oligodendrocyte transcription factor 3</i>)
OR	Odds ratio
PARP	Poli-ADP-ribosa polimerasa
pDC	Célula dendrítica plasmacitoide (del inglés, <i>plasmacytoid dendritic cell</i>)
PDCD1	Gen de muerte celular programada 1 (del inglés, <i>programmed cell death 1</i>)
PERP	<i>p53 apoptosis effector related to PMP22</i>
PHRF1	<i>PHD and ring finger domain-containing protein 1</i>
PI3K	Fosfatidil inositol 3 quinasa
PLC	Fosfolipasa C
PRDM1-ATG5	<i>PR domain containing 1, with ZNF domain-APG5 autophagy 5-like</i>
PRKCQ	Proteín quinasa C theta
PTPN22	Fosfatasa de tirosina tipo 22 no-receptor (del inglés, <i>protein tyrosin phosphatase non-receptor type 22</i>)
PXK	Quinasa de serina-treonina con dominio homólogo a <i>Phox</i> (del inglés, <i>Phox homology domain-containing serine-threonine kinase</i>)
RUNX1	<i>Runt-related transcription factor 1</i>
SLEGEN	Consorcio Internacional de Genética del Lupus Eritematoso Sistémico (del inglés, <i>Systemic Lupus Erythematosus Genetics</i>)
SNP	Polimorfismo de un sólo nucleótido (del inglés, <i>single nucleotide polymorphism</i>)
SNRPC	Polipéptido C de ribonucleoproteína nuclear pequeña (del inglés, <i>small nuclear ribonucleoprotein polypeptide C</i>)
SS	Síndrome de Sjögren
STAT	Transductor de señal y activador de la transcripción (del inglés, <i>signal transducer and activator of transcription</i>)

SV40	<i>Simian virus 40</i>
T1D	Diabetes tipo 1 (del inglés, <i>type 1 diabetes</i>)
tagSNP	SNP marcador
TCR	Receptor de células T (del inglés, <i>T-cell receptor</i>)
TDT	Análisis de desequilibrio de transmisión (del inglés, <i>transmission disequilibrium test</i>)
TGFβ	Factor de crecimiento transformante β (del inglés, <i>transforming growth factor β</i>)
Th	Linfocito T colaborador (del inglés, <i>T helper</i>)
TLR	Receptor tipo Toll (del inglés, <i>Toll-like receptor</i>)
TNF	Factor de necrosis tumoral (del inglés, <i>tumor necrosis factor</i>)
TNFAIP3	Proteína 3 inducida por el factor de necrosis tumoral α (del inglés, <i>tumor necrosis factor α-induced protein 3</i>)
TNFR	Receptor del TNF
TNFSF4/OX40L	Miembro 4 de la superfamilia de ligandos del factor de necrosis tumoral (del inglés, <i>tumor necrosis factor (ligand) superfamily, member 4</i>)
TNIP1	Proteína 1 de interacción con TNFAIP3 (del inglés, <i>TNFAIP3-interacting protein 1</i>)
TRAF	Factor asociado al receptor del factor de necrosis tumoral (del inglés, <i>tumor necrosis factor receptor-associated factor</i>)
Treg	Linfocito T regulador
TREX1	Exonucleasa 3' reparadora 1 (del inglés, <i>three prime repair exonuclease 1</i>)
TYK2	Tirosín-quinasa 2
UHRF1BP1	Proteína 1 de unión a ICBP90 (del inglés, <i>ICBP90 binding protein 1</i>)
VIH	Virus de la inmunodeficiencia humana
λ	Factor de inflación
λs	Riesgo relativo entre hermanos

INTRODUCCIÓN

Introducción

1. Lupus Eritematoso Sistémico

El lupus eritematoso sistémico (LES; [MIM] 152700) es una enfermedad autoinmune crónica de etiología compleja. Se caracteriza por una alteración del sistema inmune innato y adquirido, que lleva a la pérdida de tolerancia frente a antígenos propios y a la consecuente producción de autoanticuerpos y daño tisular. El LES es considerado una enfermedad multisistémica, ya que puede afectar a diversos órganos, como articulaciones, riñones, corazón, sistema nervioso, pulmones, sistema hematológico o piel (**Figura 1**). El curso de la enfermedad es impredecible, con periodos de crisis alternando con otros de remisión.

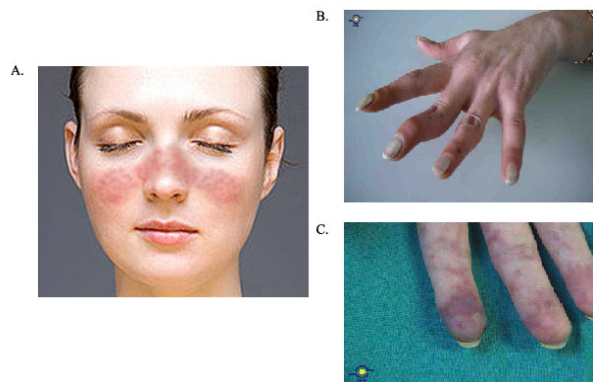


Figura 1. Algunas manifestaciones clínicas características del LES. A. Eritema malar, en forma de alas de mariposa. B. Artritis no erosiva, que se manifiesta preferentemente en las manos. C. Vasculitis.

El LES es una enfermedad que afecta con una frecuencia 9 veces mayor a mujeres que a hombres, sobre todo a mujeres en edad fértil. La prevalencia en población europea es del 0.01-0.06 %, mientras que es tres o cuatro veces más frecuente en poblaciones asiática, africana e hispano-americana (Petri M y col., 2002). Esta heterogeneidad étnica junto con la heterogeneidad genética pueden contribuir a la complejidad de sus manifestaciones clínicas.

1.1 Patogenia del LES

Los mecanismos que llevan a la aparición de la enfermedad, así como su orden temporal, no son bien conocidos. Un posible mecanismo desencadenante de la patogenia del LES sería la elevada tasa de apoptosis de linfocitos, monocitos y macrófagos en la circulación, la deficiente eliminación de los restos apoptóticos y un reconocimiento anormal de autoantígenos liberados durante este proceso. Estos autoantígenos son captados y procesados por las células presentadoras de antígeno (CPA). Éstas son, entre otras, células dendríticas (DC) y linfocitos B, cuya función es presentar los antígenos a otras células del sistema inmune mediante el complejo mayor de histocompatibilidad (MHC). En el LES, las DCs presentan autoantígenos a linfocitos T inmaduros, llevando a su activación. Las células B presentan autoantígenos a linfocitos T colaboradores (Th), en los que inducen la producción de citocinas. Las citocinas, a su vez, estimulan la proliferación de las células B y la producción de autoanticuerpos frente a los autoantígenos nucleares, citoplasmáticos y de superficie celular. Las células T y B activadas se infiltrarían en los tejidos y, junto con los autoanticuerpos y depósitos de complejos inmunes, desencadenarían procesos inflamatorios (activación del complemento, quimiotaxis) provocando daño tisular y fibrosis (Cooper GS y col., 2008) **(Figura 2).**

“asesinas naturales”. Estos dos tipos de células serían las encargadas de limitar la expansión de las células Th y B autorreactivas, por lo que la alteración de su función produce un defecto en la tolerancia periférica.

En los pacientes con LES también se ha observado una función alterada de las células Th. El desequilibrio en la producción de citocinas por linfocitos Th1/Th2 puede tener un papel importante en la patogenia del LES. Así, se ha visto una respuesta defectuosa de las células Th1, productoras de interleucina (IL) 12, IFN γ (IFN de tipo II) y linfoxina (también llamado factor de necrosis tumoral β , TNF β), mientras que tiene lugar una proliferación incrementada de los linfocitos Th2, productores de IL4, IL5 e IL13. De hecho, la IL4, además de promover la diferenciación de las células Th2, inhibe la proliferación de las células Th1. Como resultado de este desequilibrio, se produce hiperactividad de las células B con un aumento de la producción de autoanticuerpos (Mok CC y col., 2003). Recientemente, se ha descubierto que otra subpoblación de células Th, las Th17 tienen un papel accesorio, pero no por ello menos importante, en este proceso autoinmune. Su nombre se debe a que son las principales productoras de IL17A e IL17F, citocinas importantes del sistema inmune adquirido por su papel proinflamatorio. Estas citocinas ejercen su efecto en el LES facilitando la infiltración y activación de células T y B (Nalbandian A y col., 2009).

Los procesos de apoptosis o inmunidad también pueden verse alterados por cambios epigenéticos (Zhao S y col., 2009, Ballestar E y col., 2006). La epigenética hace referencia a los cambios estables y potencialmente heredables en la expresión génica que no implican cambios en la secuencia del ácido desoxirribonucleico (ADN). Las principales marcas epigenéticas son la metilación del ADN y las modificaciones de las histonas, tales como

acetilaciones, metilaciones y fosforilaciones. Éstas modulan la arquitectura de la cromatina haciéndola más o menos accesibles, lo que afecta a la actividad transcripcional. La alteración de estos procesos podría tener un papel en el LES. Por ejemplo, inhibidores de las desacetilasas de histonas reducen la expresión de múltiples genes, como *IL6*, *IL10*, *IL12* e *IFN γ* , previniendo el daño renal en ratones con un modelo de lupus (Mishra N y col., 2003). Por otra parte, la metilación de los motivos CpG del ADN produce el silenciamiento de genes codificantes de citocinas de células Th. Este efecto puede ser importante ya que se ha observado que pacientes con LES tienen hipometilado el ADN (Richardson B y col., 1990) y que inhibidores de la metilación del ADN o bien la inyección de células T demetiladas en ratones lleva a la aparición de manifestaciones de tipo lupus (Yung RL y col., 1995, Yung R y col., 2001). Estos resultados explicarían que sólo el ADN procedente de células apoptóticas, y no el ADN normal o el necrótico, pueda actuar como autoantígeno induciendo autoinmunidad, ya que el ADN apoptótico se caracteriza por presentar niveles reducidos de metilación (Wen ZK y col., 2007). En parte, estas alteraciones podrían venir explicadas por variantes en el gen de la proteína 2 de unión a CpG metilado (*MECP2*), que se ha visto significativamente asociado con LES (Sawalha AH y col., 2008).

Los miRNAs también pueden ejercer un papel en las enfermedades autoinmunes mediante su acción en los linfocitos T y B. Los miRNA son moléculas de ácido ribonucleico (ARN) no codificante de 21-24 nucleótidos que regulan la expresión de genes diana a nivel de la traducción. La fosfatasa de tirosina tipo 22 no-receptor (*PTPN22*), inhibidora de la respuesta de las células T frente a antígenos, ve reducida su expresión por el miR-181a (Li QJ y col., 2007). Por otra parte, miR-146 se une a la región 3'UTR de genes implicados en la ruta del IFN, actuando como un regulador negativo. Este miRNA inhibe la

expresión del factor asociado al receptor del factor de necrosis tumoral (*TRAF*) 6, de la quinasa 1 asociada al receptor de IL1 (*IRAK1*) (Taganov KD y col., 2006), del transductor de señal y activador de la transcripción (*STAT*) 4 y del factor regulador de IFN (*IRF*) 5 (Tang Y y col., 2009). En concordancia con estas funciones, miR-146 tiene una expresión disminuida en pacientes con LES, correlacionando negativamente con la actividad de la enfermedad y los niveles de IFN (Tang Y y col., 2009).

1.2. Clínica del LES

Según el Colegio Americano de Reumatología (ACR), un paciente se clasifica como LES si presenta al menos 4 de 11 criterios establecidos (Tan EM y col., 1982; Hochberg MC y col., 1997) (**Tabla 1**), lo que permite una gran heterogeneidad de fenotipos. Algunas de las manifestaciones incluidas en los criterios de la ACR, como fotosensibilidad, úlceras orales y artritis, son relativamente leves, mientras que otras, como alteraciones renales, convulsiones, psicosis y anormalidades hematológicas severas caracterizan una enfermedad más grave.

La producción de autoanticuerpos también está relacionada con la severidad de la enfermedad. De entre los autoanticuerpos producidos en el LES, los más característicos son los anticuerpos antinucleares (ANAs). Éstos están presentes en el 95 % de los pacientes, por lo que se incluyen como uno de los criterios de clasificación del LES. Los ANAs incluyen múltiples especificidades. Entre ellos destacan los anticuerpos anti-ADN de doble cadena, anti-histonas, anti-nucleosomas y anti-ribonucleoproteínas (anti-RNP). Algunos de estos anticuerpos suelen estar presentes antes de que se diagnostique la enfermedad y su acumulación progresiva puede llevar a la aparición de determinados

fenotipos clínicos (Arbuckle MR y col., 2003). Por ejemplo, los anti-ADN de doble cadena se asociaron con la enfermedad renal y los anticuerpos anti-Ro/SSA y anti-La/SSB están asociados con manifestaciones dermatológicas y lupus neonatal (Boh EE, 2004). Otros anticuerpos importantes son los anticuerpos antifosfolípido (aPL) (ej, anti-cardiolipina), que reconocen moléculas de la membrana celular y están relacionados con el síndrome antifosfolípido secundario en el LES, caracterizado por trombosis recurrente y abortos de repetición (Wilson WA y col., 1999).

1. Eritema Malar
2. Eritema Discoide
3. Fotosensibilidad
4. Úlceras Orales
5. Artritis
6. Serositis
 - a. Pleuritis
 - b. Pericarditis
7. Alteración Renal
 - a. Proteinuria
 - b. Cilindros Celulares
8. Alteración Neurológica
 - a. Convulsiones
 - b. Psicosis
9. Alteración Hematológica
 - a. Anemia Hemolítica
 - b. Leucopenia
 - c. Linfopenia
 - d. Trombocitopenia
10. Alteración Inmunológica
 - a. Anticuerpos anti-ADN de doble cadena
 - b. Anticuerpos anti-Smith (anti-Sm)
 - c. Anticuerpos Antifosfolípido (aPL)
 - Anticuerpos Anticardiolipina
 - Anticoagulante Lúpico
 - Falso Positivo para Sífilis
11. Anticuerpos Antinucleares (ANAs)

Tabla 1. Criterios de clasificación del LES establecidos en 1982 (Tan EM y col., 1982) y actualizados en 1997 (Hochberg MC y col., 1997).

A pesar de todos estos fenotipos que se asocian con los autoanticuerpos, hasta un tercio de la población general puede presentar algún tipo de los mismos (Teubner A y col., 2002). Pero sólo el 1 % de personas con ANAs desarrollará LES (Petri M y col., 2002). Por tanto, la mayoría de individuos con autoanticuerpos están protegidos frente a sus consecuencias patológicas. Esto sugiere que la autoinmunidad contribuye al desarrollo de lesiones tisulares, pero no es suficiente para su producción.

1.3. El LES como enfermedad compleja

Parte de la complejidad del LES radica en que es debido a la interacción de múltiples factores, tanto ambientales y genéticos, como hormonales (Tsao BP y col., 2003). La importancia del componente genético se ha puesto de manifiesto por varias observaciones. En primer lugar, el LES tiende a agruparse en familias, aunque no sigue un patrón de herencia mendeliana simple. Aún así, tiene una heredabilidad importante, de más del 60 %. Por otra parte, el riesgo relativo a padecer la enfermedad entre hermanos (λ_s) es 20-40 veces mayor que el riesgo general de la población. Además, la enfermedad es concordante entre gemelos que comparten el mismo ambiente en el 2-5 % de los dicigotos y en el 24-58 % de los monocigotos. Esta diferencia de 10 veces en la concordancia de la enfermedad entre gemelos idénticos (los que comparten todos sus genes) y gemelos-hermanos (que comparten la mitad de sus genes) sugiere que múltiples genes con un efecto modesto, compartidos por cada par de gemelos, tienen una influencia importante en la susceptibilidad al LES (Tsao BP y col., 2003; Moser KL y col., 2009; Alacon-Segovia D y col., 2005). Así que todas estas observaciones muestran la contribución de un componente genético significativo. De todos modos, como la concordancia más alta es del 58 %,

otros factores no genéticos también tienen que estar contribuyendo a la etiología del LES.

En cuanto al componente hormonal, se han propuesto los estrógenos femeninos como promotores de la enfermedad. Éstos podrían ser los responsables de la mayor prevalencia del LES en mujeres (9 mujeres por cada hombre). Por otro lado, el componente ambiental posiblemente incluye factores tales como agentes mutagénicos (luz ultravioleta), infecciones víricas (virus Epstein-Bar), la dieta y el estilo de vida (obesidad, hábito de fumar). Estos factores modularían la susceptibilidad al LES alterando los procesos de apoptosis o inmunidad, o a través de cambios epigenéticos (Zhao S y col., 2009; Ballestar E y col., 2006).

2. Influencia del componente genético en la variabilidad clínica del LES

El componente genético no sólo influencia la susceptibilidad sino la forma clínica, como se ha demostrado en estudios epidemiológicos en Estados Unidos. Éstos han observado que el riesgo a padecer manifestaciones severas de LES, así como la edad de inicio de la enfermedad, varían según el grupo étnico (hispanos, afro-americanos o caucásicos). En poblaciones no caucásicas los pacientes tienden a ser más jóvenes y a sufrir una enfermedad más grave y activa, con afectación de sistemas como el renal o el cardiovascular y una alta producción de autoanticuerpos. Esto sugiere que la genética podría influenciar la prevalencia de los distintos fenotipos (Alarcón GS y col., 2002; Alarcón GS y col., 1999).

En el mismo sentido, se sabe que existen diferencias genéticas entre poblaciones del mismo continente, lo que sugiere que también podría existir una contribución de la ascendencia genética, dentro de una misma etnia, al desarrollo de las distintas manifestaciones clínicas del LES. Así, en Europa existe un gradiente en la variabilidad genética en dirección norte-sur, que permite distinguir entre poblaciones noreuropeas (Reino Unido, Alemania, Escandinavia y Europa del este) y poblaciones sureuropeas (Italia, Grecia y España) (Novembre J y col., 2008; Seldin MF y col., 2006). Un estudio reciente ha demostrado la asociación de estas subpoblaciones de europeos con fenotipos específicos del LES. Concretamente, se ha observado una mayor prevalencia de manifestaciones mucocutáneas (fotosensibilidad y eritema discoide) en los sujetos con ascendencia del norte de Europa, mientras que en los que procedían del sur era mayor la prevalencia de autoanticuerpos y de nefritis (Chung SA y col., 2009; Richman IB y col., 2009).

Los pacientes con origen en Europa occidental presentaban más serositis y producción de autoanticuerpos en comparación con los pacientes de origen oriental. También se ha visto que los judíos askenazíes presentan menos alteraciones neurológicas del LES. Estas diferencias sugieren que incluso pequeñas diferencias en la ascendencia genética podrían influenciar el desarrollo de manifestaciones específicas (Richman IB y col., 2009). Estos estudios se realizaron con una colección de muestras caucásicas recogidas en Estados Unidos, con un gran predominio de los pacientes con origen noreuropeo, por lo que sería interesante abordar esta cuestión con muestras obtenidas en Europa y con una representación más equilibrada.

3. Estudio de la genética del LES

Como las variantes genéticas están presentes desde el nacimiento y no están afectadas por el curso de la enfermedad ni por el tratamiento, los estudios genéticos serían capaces de identificar factores de susceptibilidad y mecanismos fundamentales de la patogenia. Por lo tanto, estos estudios nos permitirían entender cómo la falta de regulación de un número de procesos inmunológicos clave predisponen al desarrollo del LES. Se ha hipotetizado que el desarrollo de la enfermedad tendría lugar cuando el contenido acumulado de alelos de susceptibilidad excede un umbral, y sería desencadenado por procesos no genéticos de naturaleza poco clara, como infecciones víricas.

Los dos tipos de estudios que se utilizan para buscar factores genéticos de susceptibilidad son los estudios de ligamiento y los estudios de asociación.

3.1. Estudios de ligamiento

Los estudios de ligamiento analizan la cosegregación de secuencias genómicas entre individuos de familias con múltiples enfermos. Se estudian marcadores (preferentemente microsatélites) repartidos por todo el genoma, de modo que si un marcador está en ligamiento con un *locus* de susceptibilidad, los miembros enfermos de la familia tendrán el mismo alelo del marcador. De esta manera se observan las regiones cromosómicas que segregan con la enfermedad y se buscan posibles genes candidatos en las mismas.

Los estudios de ligamiento son el método de elección para estudiar los fenotipos de herencia mendeliana, porque los modelos genéticos son pocos y de fácil análisis. También son una buena aproximación para encontrar variantes

raras con grandes efectos, es decir, que provocan una alta probabilidad de desarrollar la enfermedad. Sin embargo, su aplicación a los fenotipos complejos, como el LES, es más problemática porque no existe un modelo que explique adecuadamente el patrón de herencia y porque los factores genéticos de susceptibilidad para estas enfermedades suelen ser muchos y tener un efecto pequeño. De hecho, los estudios de ligamiento de genoma completo realizados hasta el momento en diversas poblaciones (europea, hispana, afro-americana y asiática) indican que la susceptibilidad al LES vendría dada por múltiples *loci* de pequeño efecto y de magnitud similar (Cantor RM y col., 2004; Gaffney PM y col., 1998; Gaffney PM y col., 2000; Gray-McGuire C y col., 2000; Johansson CM y col., 2004; Koskenmies S y col., 2004; Lindqvist AK y col., 2000; Moser KL y col., 1998; Nath SK y col., 2004; Shai R y col., 1999, Rao S y col., 2001; Olson JM y col., 2002). Sin embargo, no hay mucha concordancia entre los resultados. De los casi 60 *loci* que han demostrado evidencias de ligamiento con LES en estudios de genoma completo, sólo 17 han sido reproducidos en, al menos, dos estudios independientes: *1p36*, *1q23-24*, *2q23-33*, *2q37*, *4p15-13*, *4q32*, *5p15*, *6p22-q15*, *11q23*, *12q24*, *14q23*, *16p13*, *16q12-13*, *17q11*, *18q21*, *19q13.1* y *20p13*. En consecuencia, para encontrar alguna evidencia consistente en los estudios de ligamiento realizados, se han llevado a cabo dos meta-análisis (Lee YH y col., 2005; Forabosco P y col., 2006). Ambos confirmaron los *loci* *6p22-q15*, *16p12-q12*, *20p12-q13* y *2q31-q34*, proponiéndose como genes causantes del ligamiento de los dos primeros *loci* al *MHC* y a *CARD15/NOD2* (*caspase recruitment domain 15 / nucleotide-binding oligomerization domain containing 2*), respectivamente. De todos modos, el *MHC* no tiene un efecto tan fuerte como en otras enfermedades autoinmunes como la artritis reumatoide (AR).

La falta de reproducibilidad que se observó en los estudios de ligamiento de LES puede deberse al efecto débil de los *loci*, a la heterogeneidad genética o fenotípica de las colecciones, al tamaño muestral pequeño de los estudios, a diferencias en el poder de detección del ligamiento o en el método estadístico usado. Por otra parte, la región que identifican como ligada a la enfermedad es muy grande, y hacen falta estudios más finos que detecten el polimorfismo responsable de dicho ligamiento. Debido a estas limitaciones, hoy en día, los estudios de ligamiento han sido relegados a un segundo plano por los estudios de asociación. Éstos últimos tienen un mayor poder estadístico, permiten el uso de conjuntos de muestras más grandes, al no basarse en familias de enfermos, y pueden evaluar directamente los posibles factores de riesgo.

3.2. Estudios de asociación

Una vez identificadas las regiones que segregan con la enfermedad, el siguiente paso sería identificar el polimorfismo causal en las mismas. Esto no es fácil, ya que las regiones de ligamiento suelen ser bastante grandes (10-30 Mb) y pueden contener de 100 a 300 genes. Para ello, se han llevado a cabo estudios de clonaje posicional y de asociación. Los estudios de asociación incluyen tanto diseños caso-control, como análisis de desequilibrio de transmisión basado en familias (TDT).

En el análisis TDT se estudia si existe una transmisión preferente de un alelo o haplotipo específico desde los padres heterocigotos a los hijos enfermos, lo que proporcionaría evidencias de asociación del alelo o haplotipo analizado con la susceptibilidad a la enfermedad (Tsao BP y col., 2003). Este análisis requiere la información genotípica de tríos constituidos por el hijo afecto y sus padres. Al menos uno de los progenitores ha de ser heterocigoto para que la

familia sea informativa, lo cual es una limitación de estos estudios. El análisis estadístico se basa en la χ^2 de McNemar's que evalúa la hipótesis nula de herencia mendeliana en la que los alelos son transmitidos al 50 % de los hijos desde los padres heterocigotos frente a la hipótesis alternativa de que un alelo se transmite más frecuentemente a la descendencia afectada. Esta metodología permitió reducir el intervalo de ligamiento del cromosoma 6 al segmento cromosómico que contiene los genes MHC de clase II en un estudio de LES (Graham RR y col., 2002).

Los estudios de asociación caso-control no se basan en familias y, por tanto, no tienen en cuenta los patrones de herencia. Se basan en comparaciones de individuos enfermos no relacionados con individuos sanos de la población, también sin relación de parentesco, lo que permite obtener grandes tamaños muestrales que aumenten la potencia estadística del estudio. El método estadístico para esta comparación es sencillo, basándose en el análisis χ^2 de tablas de contingencia 2 x 2, construidas con las frecuencias de los dos alelos de un polimorfismo en casos y controles. De este modo, se dice que uno de los alelos está asociado con la enfermedad si éste tiene una frecuencia significativamente mayor en los casos en comparación con los controles, o viceversa. Una de las dificultades de los estudios de asociación es la elección del grupo control, en contraposición con los estudios de ligamiento o los TDT, que no requieren un grupo control porque estudian modelos de herencia basados en familias. Otro problema de los estudios de asociación es la estratificación poblacional, que puede dar lugar a falsos positivos cuando hay diferencias en la frecuencia de la enfermedad y del marcador entre las subpoblaciones que forman la cohorte de muestras a estudiar. Además, siguen teniendo el inconveniente de la dificultad que supone detectar polimorfismos de

susceptibilidad con un efecto pequeño, aunque este problema va siendo menor al utilizar cada vez mayores tamaños muestrales.

A pesar de estos inconvenientes, los estudios de asociación caso-control, junto con los TDT, han permitido identificar, en intervalos de ligamiento, genes candidato, implicados en procesos inmunológicos (formación de complejos inmunes, transducción de señal o ruta del IFN). Así, el *locus 1q22-24* se podría explicar por los receptores de baja afinidad de la fracción constante de la IgG (*FCGR*) *FCGR2A*, *FCGR2B* y *FCGR3A* y también por la molécula de adhesión *LY9* (Cunninghame Graham DS y col., 2008^b). Como candidato para el ligamiento del *locus 1q41-42* se propuso la poli-ADP-ribosa polimerasa (*PARP*), que media la ADP-ribosilación post-traducciona l de proteínas. El ligamiento de la región *2q23-33* podría venir explicado por *STAT1-STAT4*, que inducen la expresión de genes activados por IFN (Remmers EF y col., 2007). En el intervalo de ligamiento *16p12-q13* se descubrió la asociación con LES del gen de la integrina α -M (*ITGAM*), que participa en la inflamación al regular la adhesión y migración leucocitaria (Nath SK y col., 2008). Por último, en el *locus 2q37* se propuso al gen de muerte celular programada 1 (*PDCDI*), que participa en el mantenimiento de la tolerancia periférica. Este último *locus* ha sido de gran interés ya que fue el primer gen que se encontró asociado con LES a partir de un *locus* de ligamiento.

3.2.1. *PDCDI*

El gen *PDCDI* (también conocido como *PDI* o *CD279*) codifica para un receptor que pertenece a la familia B7/CD28 que funciona como inhibidor de la actividad de linfocitos T y B. Tras su identificación como posible factor de susceptibilidad al LES en estudios de ligamiento con familias multicaso

islandesas, suecas y noruegas (Lindqvist AK y col., 2000; Magnusson V y col., 2000), se sugirió que el polimorfismo de un solo nucleótido (SNP) PD1.3 (G/A) sería el responsable. El alelo A mostraba una mayor frecuencia en los pacientes con LES en comparación con los controles en un estudio de asociación caso-control que incluía varios grupos étnicos. Además este SNP era interesante por su posible papel funcional, ya que parecía alterar el sitio de unión para el factor de transcripción RUNX1 (Prokunina y col., 2002). Diferentes estudios posteriores han mostrado asociación de *PDCDI* con susceptibilidad al LES o a determinados fenotipos de la enfermedad (Nielsen C y col., 2004; Prokunina L y col., 2004; Sanghera DK y col., 2004; Johansson M y col., 2005; Thorburn CM y col., 2007; Velazquez-Cruz R y col., 2007) y a otras enfermedades autoinmunes, como AR (Prokunina L y col., 2004) o diabetes tipo 1 (T1D) (Nielsen C y col., 2003), confirmando la importancia de este gen en los procesos de autoinmunidad. Sin embargo, se han observado discrepancias acerca de cuál es el alelo responsable de la relación entre *PDCDI* y autoinmunidad. Así, en otro estudio de asociación en población española realizado por nuestro grupo, el alelo de susceptibilidad a la enfermedad fue el alelo G de PD1.3 en lugar del A (Ferreiros-Vidal I y col., 2004). Además, utilizando una colección de muestras de distintos países europeos se observó un gradiente geográfico de la frecuencia del alelo A de PD1.3, que aumentaba del noreste al suroeste de Europa en los controles (Ferreiros-Vidal I y col., 2007). La ausencia de un gradiente similar en los pacientes sería la explicación de los resultados discordantes en los extremos del gradiente (suecos y españoles), excluyéndose un efecto directo de este alelo en la susceptibilidad al LES. Esta idea es reforzada por la existencia de diferentes estudios de asociación en los que se ha visto que otros polimorfismos de *PDCDI* podrían predisponer al LES o a otras enfermedades o fenotipos autoinmunes (James ES y col., 2005; Newby PR y col., 2007). En concordancia con estas dos conclusiones, los estudios de

asociación de genoma completo (GWAS) de LES en población europea no mostraron ninguna señal significativa en este *locus* (Hom G y col., 2008; Graham RR y col., 2008), por lo que su asociación con la enfermedad es muy cuestionada en la actualidad.

PDCD1 se expresa en linfocitos T y B, células NK, monocitos y células mieloides, pero sólo tras su activación (Sharpe AH y col., 2002). También se ha observado en el timo, donde puede jugar un papel en el mantenimiento de la tolerancia central (Blank C y col., 2003). Los ligandos de *PDCD1* son también miembros de la familia B7: PD-L1 (B7-H1 o CD274) (Freeman GF y col., 2000) y PD-L2 (B7-DC o CD273) (Latchman Y y col., 2001). Ambos se expresan en monocitos y DCs. PD-L1 es producido además en células T y B y en una amplia variedad de tejidos no linfoides (Coyle AJ y col., 2001). *PDCD1*, al unirse con sus ligandos, atenúa la respuesta de los linfocitos activados. La unión induce la fosforilación de las tirosinas de los dominios inmunoinhibidor ITIM (*immunoreceptor tyrosine-based inhibition motif*) y conmutador ITSM (*immunoreceptor tyrosine-based switch motif*) (Leibson PJ y col., 2004). ITSM fosforilado recluta fosfatasas de tirosina, homólogas a la fosfatasa Src2 con dominios SH2, que defosforilan moléculas como Syk, la fosfatidil inositol 3 quinasa (PI3K), la fosfolipasa C (PLC), Vav, ZAP70, la proteína quinasa C theta (PRKCC) o CD3 ζ (Sharpe AH y col., 2002). En los linfocitos T activados esta defosforilación conduce a una disminución de la proliferación y producción de citocinas como IFN γ , IL10, IL4 e IL2 (Carreno BM y col., 2002); mientras que en los linfocitos B conduce a la inhibición de procesos relacionados con la proliferación y diferenciación celular, como el cambio de clase de inmunoglobulinas (Ig) (Nishimura H y col., 1998).

La importancia de *PDCD1* en la prevención de la autoinmunidad se ha demostrado en los ratones *knockout* (KO). Los ratones de distintas cepas muestran enfermedades distintas, pero todas con etiología autoinmune. Así, los ratones de la línea BALB/c-*PDCD1*^{-/-} desarrollan una cardiopatía dilatada autoinmune, mediada por anticuerpos anti-troponina-1, que inducen la estimulación crónica del flujo de Ca²⁺ en los cardiomiocitos (Okazaki T y col., 2003). Por su parte, la línea C57BL/6-*PDCD1*^{-/-} desarrolla un fenotipo más relacionado con el LES: glomerulonefritis de tipo lupus, depósito de IgG3 y del factor del complemento C3 en los glomérulos, artritis y esplenomegalia (Nishimura H y col., 1999). Por último, los ratones NOD (*non-obese diabetic*)-*PDCD1*^{-/-} desarrollan T1D más temprana que los ratones NOD *wild type*, con una penetrancia del 100 % debida a una rápida infiltración de los islotes pancreáticos por los linfocitos Th1 (Wang J y col., 2005).

Varios estudios han demostrado que *PDCD1* también es un elemento clave en la persistencia de infecciones víricas crónicas. En enfermos de SIDA, *PDCD1* está sobreexpresado en los linfocitos T CD8+ específicos del virus de inmunodeficiencia humana (VIH). Esta sobreexpresión provoca una acción antivírica ineficiente. Estos linfocitos tienen una capacidad reducida de proliferar y de producir citocinas y moléculas efectoras. Además, la expresión aumentada de *PDCD1* en los linfocitos T se correlaciona con una mayor carga viral, un aumento de apoptosis de los linfocitos T CD8+ y una disminución en el recuento de linfocitos T, características predictoras de la progresión de la enfermedad. Una demostración de la importancia de *PDCD1* en el SIDA se obtuvo al observar que el bloqueo de *PDCD1* restaura las funciones de los linfocitos T CD8+. (Day CL y col., 2006; Petrovas C y col., 2006; Trautmann L, 2006). De modo similar, se ha visto un aumento de expresión de *PDCD1* en linfocitos T CD8+, T CD4+ y Treg de pacientes con infección crónica por los

virus de las hepatitis B y C. Estos niveles elevados de expresión de *PDCD1* se correlacionaban también en estas enfermedades con la carga viral, el grado de inflamación y el daño hepático. También, en estos dos tipos de hepatitis el bloqueo de *PDCD1* da lugar a un aumento de la producción de IFN γ y de la proliferación de los linfocitos. Además, los pacientes que responden al tratamiento antiviral muestran una reducción en la expresión de *PDCD1* (Peng G y col., 2008; Golden-Mason L y col., 2008; Radziejewicz H y col., 2009; Kassel R y col., 2009). Por todo ello, se considera que *PDCD1* tiene un papel crítico en la persistencia de las infecciones víricas.

Así, *PDCD1* tiene un papel relevante en la regulación de la respuesta inmune, por lo que es un buen candidato a factor de susceptibilidad al LES. Sin embargo, los estudios de asociación más recientes no muestran una asociación clara.

3.3. Estudios de asociación de genoma completo

La mayoría de estudios de asociación que se han llevado a cabo y de los que se ha hablado hasta el momento en esta memoria se conocen genéricamente como estudios de asociación de gen candidato. Se centran en el análisis de uno o varios factores genéticos con el soporte de algún indicio sobre su posible papel en la enfermedad, ya sea por su papel funcional, como por evidencias de ligamiento del *locus* en el que se encuentran. El enfoque de los GWAS es distinto ya que se pretende evaluar la variabilidad genética a lo largo de todo el genoma, sin el soporte de una hipótesis previa.

Actualmente se conocen más de 14 millones de SNPs presentes en el genoma humano, pero no es necesario el análisis de toda esta variación genética

de forma directa. El descubrimiento de que el genoma humano está mayoritariamente incluido en bloques de desequilibrio de ligamiento (LD), regiones del genoma con variantes fuertemente correlacionadas (Internacional HapMap Consortium y col., 2007), permitió la posibilidad de explorar la variabilidad del genoma utilizando un número relativamente pequeño de SNPs. Este descubrimiento fue parejo al desarrollo de nuevas técnicas de biología molecular que permitieron analizar un gran número de SNPs en grandes colecciones de muestras de forma económica, de tal manera que se consigue una buena cobertura genética a lo largo del genoma humano. Las dos compañías que han proporcionado dichas herramientas son Affymetrix e Illumina. Ambas han comercializado *chips* para analizar hasta un millón de SNPs, siendo los más utilizados hasta el momento los de 300000 y 500000 SNPs.

Tras genotipados a gran escala como los que se realizan con estos *chips*, es necesario llevar a cabo un control de calidad a tres niveles: SNPs, muestras y genotipos. En cuanto a los SNPs, se comprueba la eficiencia de genotipado y si la frecuencia del alelo menor se encuentra en equilibrio de Hardy-Weinberg así como si está en concordancia con los datos de HapMap y con los duplicados internos. Con respecto a las muestras, se comprueba también la eficiencia de genotipado y la concordancia entre el sexo recogido y el observado. Y con los datos de genotipo se comprueba si existe algún problema de subestructura poblacional y si hay muestras que puedan tener una relación de parentesco oculta, lo que puede dar lugar a falsos positivos.

Para corregir la posible presencia de estratificación poblacional, se utilizan dos aproximaciones: el análisis de componentes principales y el control genómico. El control genómico usa un conjunto de marcadores no asociados con la enfermedad y elegidos al azar para estimar un factor de inflación, λ . Este

factor es una medida basada en el test χ^2 que refleja si la distribución de los genotipos se ajusta a la hipótesis nula de ausencia de estratificación poblacional ($\lambda = 1$). Un valor mayor de 1 indica un incremento del estadístico de asociación debido a estratificación poblacional (Tiwari HK y col., 2008). Este valor se utiliza como factor de corrección para todos los resultados de asociación en el estudio. Sin embargo, algunos SNPs difieren más que otros entre las distintas poblaciones, por lo que el ajuste uniforme que aplica el control genómico puede no ser óptimo. El segundo método para corregir la estratificación poblacional se basa en el análisis de componentes principales. Este análisis reduce todos los genotipos a un menor número de variables continuas no correlacionadas, llamadas componentes principales, que reflejan la máxima variabilidad genética posible. Si existen subpoblaciones en el estudio, los ejes de los componentes principales suelen tener una interpretación geográfica, por lo que fenotipos y genotipos se pueden ajustar según la variabilidad atribuida a la ascendencia, para llevar a cabo el estudio de asociación sin interferencia de estructura poblacional (Price AL y col., 2006).

Una de las ventajas de los GWAS es que no requieren de hipótesis *a priori*. Se pueden considerar como estudios generadores de hipótesis, ya que se analiza todo el genoma sin tener en cuenta el posible papel en la enfermedad o la posible funcionalidad de los SNPs. Esto ha permitido de forma efectiva descubrir genes de susceptibilidad que no se esperaba que tuviesen un papel en la enfermedad. Sin embargo, dado al alto número de hipótesis que son analizadas, se ha establecido un umbral de significación mucho más estricto ($P < 5 \times 10^{-8}$) que el clásico de $P < 0.05$. Este umbral reduce la tasa de falsos positivos lo que hace que los resultados de asociación de los GWAS sean altamente reproducibles. Pero se necesita un tamaño muestral grande para detectar incluso las asociaciones más fuertes (**Tabla 2**).

La potencia estadística de los GWAS, la probabilidad de que el GWAS encuentre una asociación significativa, no puede ser calculada de forma analítica por la existencia de LD entre los SNPs. Por ello se utilizan simulaciones. El resultado depende de los SNPs incluidos en el *chip*, del tamaño del efecto a detectar y del tamaño muestral disponible. En general, los GWAS tienen una buena potencia ($> 80 \%$) para detectar efectos fuertes (*Odds Ratio* (OR) > 1.5), pero ésta se ve muy reducida ($< 70 \%$) en la identificación de efectos más modestos (OR > 1.2).

Los GWAS tienen algunas limitaciones. La primera es que los SNPs funcionales pueden no estar en LD con ninguno de los marcadores genotipados, por lo que no serán detectados. Esto es especialmente posible para las variantes raras ($< 5 \%$), que no están bien recogidas en los *chips* utilizados. Por otra parte, los GWAS no pueden evaluar otro tipo de variantes como los polimorfismos de inserción/delección (indel) o las inversiones. Sin embargo, las últimas plataformas de genotipado ya incluyen herramientas para la detección de los polimorfismos conocidos como CNVs o variaciones en el número de copias.

3.3.1. *GWAS en LES*

En total, se han publicado 5 GWAS en LES, cuatro de ellos en población europea y uno en asiáticos. Estos estudios además de identificar un gran número de genes o *loci* de susceptibilidad, han confirmado de forma consistente la asociación con LES de factores genéticos ya conocidos, como el antígeno leucocitario humano (*HLA*), *IRF5*, *FCGR2A*, *PTPN22* y *STAT4* (Hom G y col., 2008, SLEGEN y col., 2008, Kozyrev SV y col., 2008; Graham RR y col., 2008; Han JW y col., 2009) (**Tabla 2**).

		SNP	Locus	Chr	Alelo Mayor/ Alelo Menor ^a	<i>P</i> GWA < 5x10 ⁻⁸	<i>P</i> meta- análisis
GWAS europeos	SLEGEn y col., 2008 (N=6728)	rs2476601	<i>PTPN22</i>	1p13	G/A		-
		rs1801274	<i>FCGR2A</i>	1q23	C/T		-
		rs10798269	-	1q25.1	C/T		1.11x10 ⁻⁷
		rs7574865	<i>STAT4</i>	2q32	G/T	Sí	-
		rs7601754	<i>STAT4</i>	2q32	T/C		-
		rs6445975	<i>PXK</i>	3p14.3	A/C		7.10x10 ⁻⁹
		rs3131379	<i>HLA</i>	6p21.33	G/A	Sí	1.71x10 ⁻⁵²
		rs12537284	<i>IRF5/TNPO3</i>	7q32	G/A		3.61x10 ⁻¹⁹
		rs4963128	<i>KIAA1542</i>	11p15.5	C/T		3.00x10 ⁻¹⁰
		rs9888739	<i>ITGAM</i>	16p11.2	C/T		1.61x10 ⁻²³
	Hom G y col., 2008 (N=6301)	rs7574865	<i>STAT4</i>	2q32.3	G/T	Sí	9x10 ⁻¹⁴
		rs2187668	<i>HLA</i>	6p21	G/A	Sí	3x10 ⁻²¹
		rs10488631	<i>IRF5/TNPO3</i>	7q32	T/C	Sí	2x10 ⁻¹¹
		rs13277113	<i>C8orf13-BLK</i>	8p21.3	G/A		1x10 ⁻¹⁰
		rs11574637	<i>ITGAM-ITGAX</i>	16p11.2	T/C		3x10 ⁻¹¹
	Graham RR y col., 2008 (N=2586; y 740 trios)	rs5029939	<i>TNFAIP3</i>	6q23.3	C/G	Sí	2.89x10 ⁻¹²
		rs3821236	<i>STAT4</i>	2q32.3	G/A		8.49x10 ⁻¹¹
		rs2618476	<i>BLK</i>	8p23.1	T/C		1.70x10 ⁻⁸
			<i>HLA</i>	6p21		Sí	-
			<i>IRF5/TNPO3</i>	7q32		Sí	-
	Kozyrev SV y col., 2008 (N=3971)	rs10516487	<i>BANK1</i>	4q24	G/A		3.74x10 ⁻¹⁰
GWAS asiáticos	Han JW y col., 2009 (N=12454)	rs2205960	<i>TNFSF4</i>	1q25.1	C/A		2.53x10 ⁻³²
		rs463426	<i>HIC2-UBE2L3</i>	22q11.21	A/G		1.48x10 ⁻¹⁶
		rs13385731	<i>RASGRP3</i>	2p22.3	A/G		1.25x10 ⁻¹⁵
		rs7574865	<i>STAT4</i>	2q32.3	C/A	Sí	5.17x10 ⁻⁴²
		rs10036748	<i>TNIP1</i>	5q33.1	A/G		1.67x10 ⁻⁹
			<i>HLA</i>	6p21		Sí	-
		rs548234	<i>PRDM1-ATG5</i>	6q21	A/G		5.18x10 ⁻¹²
		rs2230926	<i>TNFAIP3</i>	6q23.3	A/C		1.37x10 ⁻¹⁷
		rs4917014	<i>IKZF1</i>	7p12.2	A/C		2.75x10 ⁻²³
		rs1167796	Múltiples genes	7q11.23	G/A		2.12x10 ⁻⁸
		rs4728142	<i>IRF5</i>	7q32.1	G/A		8.14x10 ⁻¹⁹
		rs7812879	<i>BLK</i>	8p23.1	G/A	Sí	2.09x10 ⁻²⁴
		rs1913517	<i>LRRC18-WDFY4</i>	10q11.22	G/A		7.22x10 ⁻¹²
		rs4639966	Múltiples genes	11q23.3	A/G		1.25x10 ⁻¹⁶
		rs6590330	<i>ETSI</i>	11q24.3	G/A		1.77x10 ⁻²⁵
		rs1385374	<i>SLC15A4</i>	12q24.32	G/A		1.77x10 ⁻¹¹
		rs7197475	Múltiples genes	16p11.2	G/A		2.77x10 ⁻⁸

Tabla 2. (Pie de tabla en el reverso).

Tabla 2. Resumen de los hallazgos principales de los GWAS en LES. Para cada estudio se detallan los SNPs asociados, el *locus* en el que se encuentran, la posición en el cromosoma, los alelos, si la asociación alcanzó el nivel de significación requerido para los GWAS y el valor de *P* en el análisis conjunto de todos los resultados obtenidos en el mismo estudio. ^a En negrita, el alelo de riesgo.

El Consorcio Internacional de Genética del LES (SLEGEN) llevó a cabo un estudio de 317501 SNPs en 2 colecciones de muestras que sumaban un total de 720 mujeres con LES y 2337 controles. Los SNPs más consistentemente asociados fueron estudiados en otros 2 conjuntos de muestras independientes que consistían en un total de 1846 mujeres con LES y 1825 controles (SLEGEN y col., 2008). En el análisis combinado, encontraron asociación con 3 SNPs en *ITGAM* (rs9888739, rs1143678 y rs4548893), el SNP rs4963128 en *KIAA1542* y rs6445975 en la quinasa de serina-treonina con dominio homólogo a *Phox* (*PXK*). Además, el SNP rs10798269 en *Iq25.1*, que se encuentra lejos de cualquier gen conocido, fue resaltado como muy probablemente asociado (**Tabla 2**).

Un segundo estudio analizó 502033 SNPs en 1311 casos con LES y 3340 controles de población europeo-americana, realizando posteriormente una fase de confirmación de los resultados más destacables en 793 casos con LES y 857 controles de población sueca (Hom G y col., 2008). Sus resultados mostraron también la asociación significativa de polimorfismos en la región cromosómica que contiene el gen *ITGAM*. Además, vieron asociación de un *locus* que hasta ahora no había sido relacionado con LES: la región promotora de los genes de la quinasa de linfocitos B (*BLK*) y de *C8orf13*, que comparten la misma región promotora aunque se transcriben en dirección opuesta (Hom G y col., 2008) (**Figura 3**).

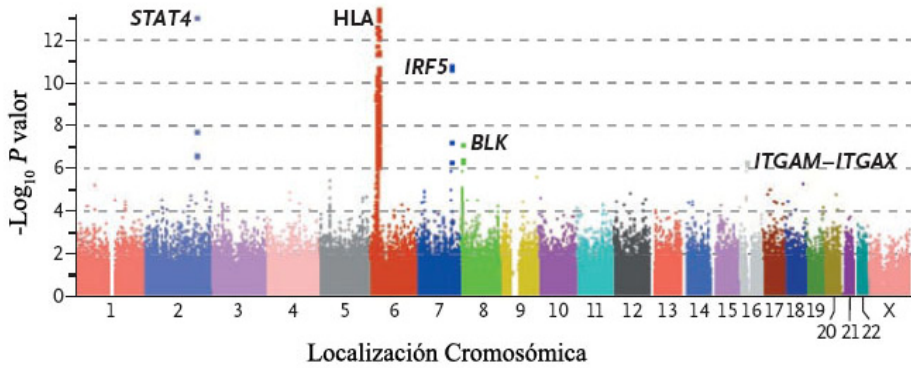


Figura 3. Diagrama de Manhattan representando por puntos todos los SNPs estudiados en el GWAS realizado por Hom y col (Hom G y col., 2008). El eje de abscisas muestra la posición en los cromosomas que están identificados por diferentes colores. En el eje de ordenadas se representa el $-\log_{10}$ del valor P correspondiente a cada SNP.

Otro de los GWAS publicados en 2008, incluyó el estudio de 311238 SNPs en 431 pacientes y 2155 controles de ascendencia europea. En este estudio se encontró otro *locus* que pasaba el umbral de significación de genoma completo: la proteína 3 inducida por $\text{TNF}\alpha$ (*TNFAIP3*). Este *locus* ya se había visto asociado con AR previamente (Graham RR y col., 2008) (**Figura 4**). Además, los 134 *loci* con menor P se analizaron también en 740 tríos y el meta-análisis de ambas fases mostró asociación de *BLK* e *ITGAM* (**Tabla 2**).

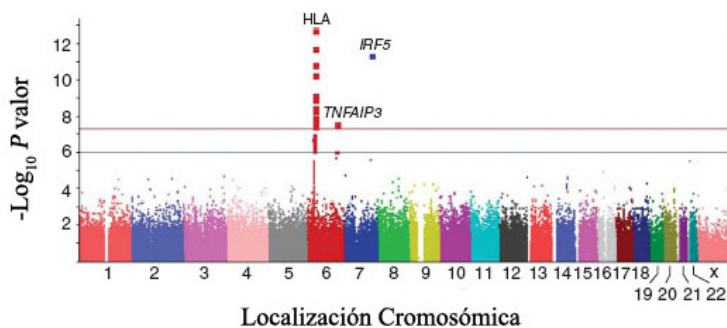


Figura 4. Diagrama de Manhattan representando los resultados del GWAS de LES realizado por Graham y col (Graham RR y col., 2008). Dos líneas punteadas horizontales indican los niveles de $P = 10^{-6}$ y $P = 5 \times 10^{-8}$.

Por último, en otro estudio de todo el genoma pero de menor densidad, se genotiparon 85042 SNPs en una cohorte sueca de 279 casos y 515 controles. Como consecuencia de este GWAS se encontró asociación de la proteína estructural de células B con repeticiones de anquirina (*BANK1*) que alcanzó el valor de $P < 5 \times 10^{-8}$ cuando se estudió en 4 cohortes adicionales de población escandinava, argentina, alemana, italiana y española (2003 pacientes y 1968 controles) (Kozyrev SV y col., 2008) (**Tabla 2**).

También se ha realizado un GWAS en población china (Han JW y col., 2009) (**Figura 5**). En este estudio se analizaron 493955 SNPs en 1047 casos y 1205 controles y las asociaciones significativas fueron evaluadas en dos colecciones de muestras adicionales que englobaban 3152 casos y 7050 controles. Como resultado, se identificaron 8 *loci* de susceptibilidad nuevos (*RASGRP3*, *IKZF1*, *ETS1*, *SLC15A4*, *7q11.23*, *10q11.22*, *11q23.3* y *16p11.2*). Además, al igual que en población europea, también observaron asociación con LES de los *loci* de la proteína 1 de interacción con TNFAIP3 (*TNIP1*) (Gateva

V y col., 2009), *PRDMI-ATG5* (*PR domain containing 1, with ZNF domain-APG5 autophagy 5-like*) (SLEGEN y col., 2008) e *HIC2-UBE2L3* (*hypermethylated in cancer 2-ubiquitin-conjugating enzyme E2L 3*) (SLEGEN y col., 2008), y confirmaron otros 5 *loci* ya más consolidados: el miembro 4 de la superfamilia de ligandos del TNF (*TNFSF4*) (Cunninghame Graham DS y col., 2008^a), *STAT4* (Hom G y col., 2008; Graham RR y col., 2008), *TNFAIP3* (Graham RR y col., 2008), *IRF5* (SLEGEN y col., 2008; Hom G y col., 2008; 2008; Graham RR y col., 2007^a) y *BLK* (Hom G y col., 2008) (**Tabla 2**).

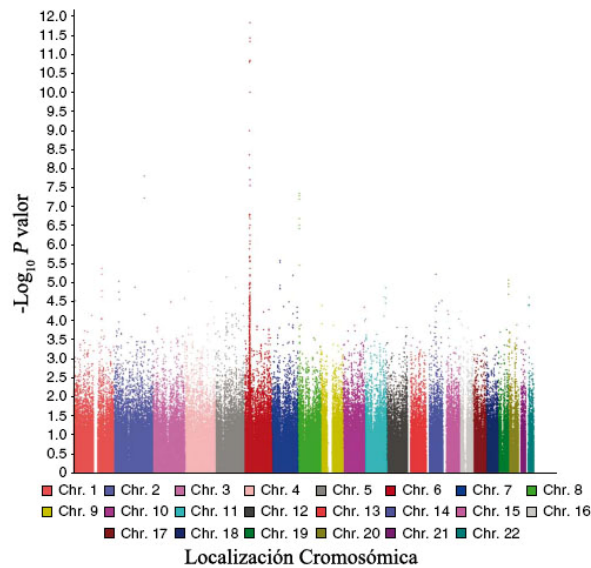


Figura 5. Diagrama de Manhattan representando los resultados del GWAS de LES en chinos realizado por Han y col (Han JW y col., 2009).

Dos herramientas complementarias tras los GWAS han sido la imputación genotípica y el meta-análisis. La imputación permite evaluar la asociación de marcadores genéticos que no han sido directamente genotipados. Estos genotipos se pueden asignar conociendo los patrones de LD en un

conjunto de muestras de referencia, como son las muestras de HapMap. La imputación también es muy útil para combinar resultados de estudios que utilizan diferentes plataformas de genotipado. La combinación de datos de dos o más estudios se denomina meta-análisis. Su objetivo es tratar de detectar asociaciones que se escapen en los GWAS individuales por tener un efecto modesto. Además, los meta-análisis también permiten comprobar la consistencia y credibilidad de las asociaciones, sirviendo como un método de confirmación. Con este objetivo, se ha publicado un meta-análisis que encontró 4 *loci* significativos según el criterio de los GWAS: *IRAK1*, *TNFSF4*, *UBE2L3* y *FCGR2A* (Graham RR y col., 2009). Tres de estos *loci* ya eran conocidos por otros estudios previos.

Otra aproximación que se ha seguido para descubrir más factores de susceptibilidad ha sido el análisis de SNPs con alguna evidencia de asociación en un GWAS en más muestras para ver si se confirman. Un estudio de este tipo (Gateva V y col., 2009) intentó identificar nuevos *loci* de riesgo entre los 2466 que mostraban una $P < 0.05$ en el GWAS de Hom y col. También incluyeron en el análisis 505 SNPs de 23 *loci* previamente identificados como asociados con LES y 42 SNPs implicados en otras enfermedades autoinmunes. El genotipado se llevó a cabo en colecciones de Estados Unidos (1129 casos y 2991 controles) y Suecia (834 casos y 1338 controles) independientes a las utilizadas en el GWAS. Se confirmaron las asociaciones previas de *TNFAIP3*, *BLK*, *ITGAM*, *KIAA1542* y *TNFSF4*, y además, se identificaron 5 nuevos *loci* de susceptibilidad al LES con un valor de P menor al umbral de significación de los GWAS: *TNIP1*, *PRDM1-ATG5*, *JAZF1* (*JAZF* zinc finger 1), proteína 1 de unión a ICBP90 (*UHRF1BP1*) e *IL10*.

A pesar de la gran cantidad de factores genéticos de susceptibilidad al LES identificados en los GWAS, hay que tener precaución en su interpretación ya que, debido al alto número de análisis que implican estos estudios, la probabilidad de encontrar falsos positivos es alta. Este inconveniente ya se intenta corregir estableciendo un estricto umbral de significación, pero además es necesario que estos resultados sean confirmados en conjuntos de muestras independientes y por laboratorios independientes, sobre todo por la dificultad de encontrar verdaderas asociaciones para las enfermedades complejas, especialmente cuando los efectos son modestos. Así, los distintos factores genéticos detectados no tienen el mismo grado de certidumbre. Hay asociaciones definitivas y consistentes, debido a que ya han sido identificadas en 2 o más estudios independientes y en los GWAS (como las de *HLA*, *IRF5*, *ITGAM*, *STAT4*, *C8orf13-BLK* y *TNFAIP3*). Otras asociaciones han sido identificadas sólo en uno de estos GWAS por lo que constituyen señales sólidas (como *BANK1*, *PXK*, *KIAA1542* y *Iq25.1*). Por otra parte, asociaciones no detectadas en GWAS pero sí en grandes estudios de asociación de genes candidato pueden considerarse como probables (como *MECP2* y *LY9*), mientras que otras pueden ser controvertidas porque los estudios muestran resultados diferentes (como la tirosín-quinasa 2, *TYK2*). Así que, la replicación de estos resultados es necesaria e importante porque proporciona una evidencia de la reproducibilidad de la asociación de un polimorfismo con la enfermedad.

3.4. Factores genéticos de LES confirmados

3.4.1. *HLA*

Los primeros genes que fueron descubiertos como factores de riesgo para el LES fueron los que codifican para el *HLA*. Las moléculas del HLA están

implicadas en el reconocimiento y la respuesta inmune, de ahí que muchos estudios genéticos de enfermedades autoinmunes se centren en ellos. El problema que presenta su estudio es que la región en la que se encuentran es de gran complejidad genética y con un alto LD, lo que dificulta la identificación de las variantes causales y también la identificación de otros posibles genes de susceptibilidad al LES.

El estudio que más contribuyó a definir la asociación del *MHC* con LES incluyó unos 50 microsatélites analizados en 334 familias multicaso. En este análisis quedó clara la importancia de los haplotipos que incluyen *HLA-DRB1* y *DQB1*, concretamente los correspondientes al *HLA-DR2* (*DRB1*1501-DQB1*0602*), *HLA-DR3* (*DRB1*0301-DQB1*0201*) y, con una menor importancia, al *HLA-DR8* (*DRB1*0801-DQB1*0402*) (Graham RR y col., 2002). Este estudio delimitó una región de 1 Mb, que engloba la región del *MHC* de clase II y III, pero no fue posible precisar más debido al alto LD de la zona. Los haplotipos portadores del *HLA-DR3* confieren mayor riesgo que los portadores del *HLA-DR2* y los individuos homocigotos para DR3 o heterocigotos DR3/DR2 muestran el riesgo más alto. Además, estos haplotipos se han asociado con la producción de autoanticuerpos característicos del LES (Graham RR y col., 2007^b; Sebastiani GD y col., 2009). Concretamente, el haplotipo DR2 se asocia con la producción de anti-Sm y anti-ADN de doble cadena, mientras que DR3 está relacionado con mayores niveles de anti-Ro y anti-La. Los GWAS publicados recientemente (SLEGEM y col., 2008; Hom G y col., 2008; Graham RR y col., 2008) han confirmado que los marcadores localizados en la región del *HLA* de clase II son los más fuertemente asociados con la enfermedad, aunque su contribución al LES no es mucho mayor que la de otros factores genéticos encontrados (**Figura 6**), por lo que el efecto del HLA no es comparable con el observado en otras enfermedades autoinmunes como

AR o T1D (Fernando MM y col., 2008). Tampoco está clara la asociación del *HLA* con LES en grupos étnicos no caucásicos (Graham RR y col., 2007^b; Fernando MM y col., 2007). Por ejemplo, en población latinoamericana ha sido un factor de susceptibilidad inconsistente, aunque es posible que esta inconsistencia sea debida a la heterogeneidad genética y el pequeño tamaño muestral de los estudios. De hecho, en un meta-análisis, con mayor potencia estadística, los alelos *HLA-DR2* y *-DR3*, así como el haplotipo *DRB1*0301-DQB1*0201* conferían un aumento significativo del riesgo de LES en esta población (Castaño-Rodríguez N y col., 2008).

Los genes *HLA-DR* y *HLA-DQ* tienen un papel muy importante en los procesos inmunes ya que las proteínas que codifican se expresan en CPAs y su función es presentar péptidos extraños a los receptores de células T (TCR) de las células Th CD4+. Estas células estimuladas inducen la producción por las células B de anticuerpos específicos frente al péptido, iniciando la respuesta inmune adquirida. Sin embargo, no se conoce cuál es la participación específica de los alelos *HLA* en la susceptibilidad al LES.

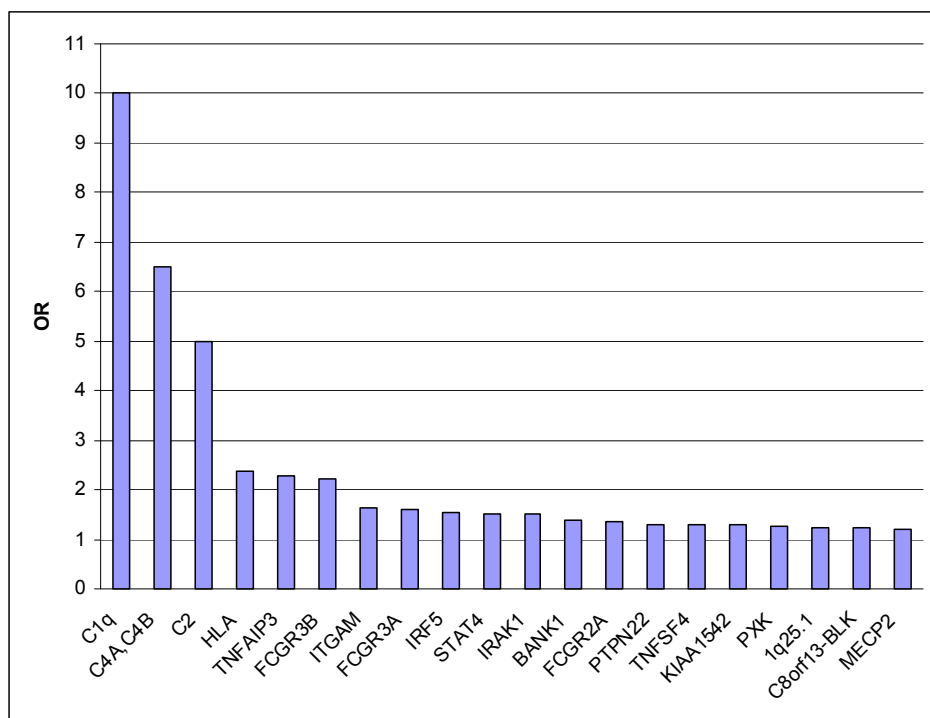


Figura 6. Comparación del efecto de los genes asociados al LES. Como medida del tamaño del efecto se usaron las ORs (Moser KL y col., 2009).

3.4.2. Factores del Complemento

Las deficiencias en los componentes del complemento C2, C4A y C4B, cuyos genes están localizados en la región del *HLA* de clase III, y C1q, con el gen localizado en *1p34.1-1p36.3*, muestran una asociación muy fuerte con LES (**Figura 6**) (Sullivan KE y col., 1994; Fielder AH y col., 1983; Reveille JD y col., 1985; Botto M y col., 2002). En concreto, la deficiencia del factor C4, bien C4A o C4B, es una de las CNVs asociadas con LES mejor definida. La deficiencia completa del gen (cuatro alelos nulos) es muy rara pero está fuertemente asociada con el LES, ya que el 75 % de individuos con esta

deficiencia sufren la enfermedad (Manderson AP y col., 2004). El alelo nulo de *C4A* está asociado en casi todas las poblaciones estudiadas. Sin embargo, esta deficiencia se encuentra en el mismo haplotipo del *HLA-DR3*, también asociado a LES, por lo que todavía no está claro si la asociación con LES de estos dos factores genéticos es o no independiente (Kelly JA y col., 2002). Por otra parte, la deficiencia de C2, que es la más frecuente en la población general, se asocia con LES en el 33 % de los individuos (Kelly JA y col., 2002). Mientras que la deficiencia del complemento que causa una mayor incidencia de LES es la de C1q (**Figura 6**). Más del 92 % de los sujetos deficientes en C1q desarrollan LES, normalmente con un fenotipo severo (Kelly JA y col., 2002). A pesar de la fuerte penetrancia de las deficiencias de los factores del complemento, éstas son variantes raras, por lo que la fracción de la susceptibilidad al LES atribuible a estos factores es muy pequeña. Por esta razón, los factores del complemento no son analizados en los pacientes con LES de forma sistemática.

Los componentes del complemento son fundamentales en la respuesta inmunitaria y en la defensa frente a agentes extraños. Están implicados en una cascada bioquímica cuyas funciones son potenciar la respuesta inflamatoria, facilitar la fagocitosis y dirigir la lisis de células infectadas. Concretamente, C1q sería el responsable de iniciar la cascada al unirse a la fracción Fc de complejos inmunes polimerizados, causando la proteólisis de los factores C2 y C4 en sus subunidades A y B. C4A favorece la degranulación de células cebadas, liberando histamina, que provoca inflamación. Mientras que C4B forma un complejo con C2A que continúa la cascada para la activación del resto de componentes del complemento. Este sistema también es importante en el LES por su papel en el aclaramiento de restos apoptóticos mediante opsonización. Por tanto, la falta de estos componentes podría llevar a una defectuosa opsonización, transporte de complejos inmunes y eliminación de

dichos restos celulares, causando inflamación y daño tisular (Kelly JA y col., 2002).

3.4.3. *FCGRs*

Los estudios de ligamiento proporcionaron evidencias de la relación con LES del *locus* del cromosoma 1 donde se encuentran los genes que codifican para los receptores Fcγ de baja afinidad (Cantor RM y col., 2004). Múltiples miembros de esta familia de inmunoglobulinas se han asociado con LES, aunque muchas de las publicaciones han obtenido resultados inconsistentes. La evidencia más clara corresponde al polimorfismo no sinónimo de *FCGR2A* (rs1801274, H131R), que altera la afinidad del receptor por subclases específicas de IgG, afectando al procesamiento de complejos inmunes (van der Pol W y col., 1998). Se han publicado más de 20 estudios de este SNP en distintos grupos étnicos, con resultados inconsistentes, aunque un meta-análisis ha mostrado un efecto moderado de susceptibilidad (Karassa FB y col., 2003). Además, el GWAS del SLEGEN mostró asociación de este SNP con $P = 6.8 \times 10^{-7}$ (SLEGEN y col., 2008). El polimorfismo no sinónimo de *FCGR3A* (F176V) que también modifica la afinidad del receptor (van der Pol W y col., 1998) ha mostrado igualmente asociación a LES en algunos estudios. Un meta-análisis mostró que está asociado pero sólo en los pacientes con nefritis lúpica (Karassa FB y col., 2003). Estos dos genes, *FCGR2A* y *FCGR3A*, están separados por tan sólo 35 kb, pero las dos asociaciones parecen independientes. *FCGR3A* se expresa en células NK, macrófagos y algunos monocitos, mientras que *FCGR2A* se expresa más ampliamente en las células hematopoyéticas. Se piensa que la baja afinidad de los alelos de riesgo por los complejos inmunes provoca una deficiencia en su eliminación, aumentando la probabilidad de una respuesta autoinmune persistente que contribuiría al desarrollo de LES.

FCGR3B, que se expresa mayoritariamente en neutrófilos, es otro gen de la familia de receptores Fc γ de baja afinidad que está implicado en la susceptibilidad al LES. Este gen presenta un polimorfismo de tipo CNV, observándose que el bajo número de copias está asociado con riesgo de LES (Fanciulli M y col., 2007; Willcocks LC y col., 2008). Además, el número de copias se ha correlacionado con la expresión del gen que, cuando es baja, implica deficiencias en la adherencia e incorporación de los complejos inmunes por los neutrófilos. Este comportamiento es característico del LES (Willcocks LC y col., 2008). *FCGR2B* es otro gen de la familia de receptores que ha mostrado asociación con LES en algún estudio (Su K y col., 2004).

La función de los Fc γ Rs consiste en unirse a complejos anticuerpo-antígeno, por la región Fc del anticuerpo, de modo que las células que expresan estos receptores pueden identificar en el LES los complejos autoinmunes para su eliminación mediante fagocitosis. Por tanto, la participación de los *FCGRs* así como del sistema del complemento en la susceptibilidad al LES indica la importancia que tiene la deficiente eliminación de complejos inmunes en el desarrollo del LES.

3.4.4. *TREX1*

Mediante secuenciación de la exonucleasa 3' reparadora 1 (*TREX1*), se han encontrado tres tipos de variantes raras en 9 de 417 pacientes caucásicos con LES, que están ausentes en 1712 controles (Lee-Kirsch MA y col., 2007). Estas variantes consisten en 2 polimorfismos que provocan un cambio en la pauta de lectura del gen, generando proteínas truncadas, SNPs no sinónimos (nsSNPs), que modifican la secuencia aminoacídica aunque no afectan al dominio catalítico, y una variante en la región no traducida 3' (3'UTR) del gen.

Mediante análisis funcionales, se observó que las proteínas truncadas veían alterada su localización celular, por lo que probablemente TREX1 no podría llevar a cabo su función.

El interés por estudiar *TREX1* en el LES se debe a varias observaciones que podrían recordar a características de esta enfermedad. Así, ya se habían identificado mutaciones en este gen relacionadas con una encefalopatía caracterizada por niveles altos de IFN α en el fluido cerebroespinal. Este fenotipo es similar al que ocurre en una forma poco frecuente del LES, denominada lupus sabañón. Además, tanto los pacientes que padecen la encefalopatía como los que presentan esta variante de LES producen ANAs. En ratones también se observó un posible papel de *TREX1* en autoinmunidad, ya que la deficiencia de este gen provoca miocarditis inflamatoria autoinmune no infecciosa (Lee-Kirsch MA y col., 2007).

TREX1 codifica para una exonucleasa 3' que metaboliza ADN de cadena sencilla. TREX1 forma parte del complejo SET, que se localiza en el retículo endoplasmático. Este complejo es traslocado al núcleo en respuesta a estrés oxidativo inducido por grancima A. El estrés oxidativo produce apoptosis independiente de caspasas, durante la cual TREX1 degrada rápidamente ADN dañado de cadena sencilla (Chowdhury D y col., 2006). TREX1 también degrada ADN de cadena sencilla derivado de virus y de retroelementos endógenos. Esta función produce una regulación negativa de la respuesta de IFN tipo I que inducirían estos ácidos nucleicos (Stetson DB y col., 2008). Las infecciones víricas se han considerado factores ambientales que podrían desencadenar el desarrollo del LES por medio de la activación de la respuesta de IFN tipo I. Por lo tanto, la posible función de TREX1 en el reconocimiento y degradación de ADN viral de cadena sencilla en el citoplasma, hacen de este

gen un buen candidato para tener un papel en dicha respuesta característica del LES.

3.4.5. *IRF5*

IRF5 codifica para un factor de transcripción que regula la expresión de genes de respuesta a IFN de tipo I. Este gen ha sido uno de los factores asociados de forma más consistente con el LES, con $OR \geq 1.5$ (**Figura 6**) (Sigurdsson S y col., 2005; Graham RR y col., 2007^a; Cunningham Graham DS y col., 2007^b; Kozyrev SV y col., 2007; Ferreiro-Neira I y col., 2007; Sigurdsson S y col., 2008^a; SLEGEN y col., 2008). En población europea, se han identificado múltiples polimorfismos de susceptibilidad, incluyendo: un SNP que crea un sitio de *splicing* en el exón 1B (rs2004640); un SNP que crea un sitio de poliadenilación en 3'UTR (rs10954213), resultando en una variante transcripcional más corta y de mayor estabilidad; un indel de 30 pares de bases en el exón 6, que afecta a un dominio de interacción de la proteína; y un indel de 5 pb en el promotor, que modifica un sitio de unión al factor de transcripción SP1. La asociación de estos SNPs es compleja, observándose un efecto protector y otro de susceptibilidad, causados por polimorfismos independientes y observándose también un posible efecto epistático entre polimorfismos funcionales (Ferreiro-Neira I y col., 2007). También se ha confirmado la asociación de *IRF5* con LES en otros grupos étnicos (Shin HD y col., 2008; Kawasaki A y col., 2008^b; Han JW y col., 2009; Kelly JA y col., 2008), aunque no está claro si las señales son similares a las identificadas en caucásicos.

Se ha sugerido que *IRF5* induciría la producción de IFN de tipo I en pDCs además de regular la expresión de genes de respuesta a IFN. Así, se ha observado que pacientes con LES portadores de haplotipos de riesgo de *IRF5*

tienen niveles más altos de IFN α (Niewold TB y col., 2008). A su vez, la producción de IRF5 es inducida por IFN α , por lo que un bucle de retroalimentación positiva puede ser parte del mecanismo que predispone al LES.

3.4.6. *PTPN22*

Alteraciones en la señalización del TCR pueden incrementar el riesgo de autoinmunidad. A este nivel actúa la fosfatasa de tirosina linfocítica (LYP), codificada por el gen *PTPN22*. Su función es inhibir la cascada de señalización del TCR. El polimorfismo R620W (rs2476601) en el exón 14 de este gen está implicado en múltiples enfermedades autoinmunes (Chung SA y col., 2007): T1D, AR, LES y enfermedad tiroidea autoinmune entre otras. Su efecto en LES parece más fuerte en los casos familiares que en los esporádicos (Kaufman KM y col., 2006). El alelo de riesgo 620W impide la interacción de LYP con la tirosín-quinasa CSK (*C-terminal Src tyrosine kinase*), que es una supresora de las quinasas que median la activación del TCR. Esta evidencia llevó a sugerir que el alelo de riesgo implicaría una pérdida de función de LYP. Por lo tanto, este alelo llevaría a una disminución en la regulación negativa de los linfocitos T y su respuesta exagerada facilitaría la autoinmunidad. Sin embargo, experimentalmente se ha observado lo opuesto. El alelo de riesgo implica una ganancia de función de LYP, resultando en una mayor inhibición de la señalización del TCR (Vang T y col., 2005). Se sugiere que esta hiposensibilidad linfocitaria produce fallos en la eliminación de células T autorreactivas en el timo, causando enfermedades autoinmunes (Rieck M y col., 2007). Un trabajo reciente ha identificado otro polimorfismo de *PTPN22* asociado al LES, R263Q, que también afectaría a la función fosfatasa. El alelo menor confiere protección frente a LES y provoca una reducción de la actividad

fosfatasa de LYP. De esta forma, se refuerza la hipótesis de que la inhibición de la actividad de LYP es un objetivo terapéutico importante (Orrú V y col., 2009).

3.4.7. *STAT4*

STAT4 se localiza en la región de ligamiento *2q* descrita tanto en LES (Cantor RM y col., 2004) como en AR (Jawaheer D y col., 2003). Estudios en población europeo-americana, identificaron este factor genético como responsable de dicha señal en las dos enfermedades (Remmers EF y col., 2007; Taylor KE y col., 2008). El SNP más asociado, rs7574865, tiene un efecto mayor en LES que en AR, con OR de 1.55 y 1.27, respectivamente. Varios estudios han confirmado esta asociación con el LES (SLEGEN y col., 2008; Graham RR y col., 2008; Palomino-Morales RJ y col., 2008; Kobayashi S y col., 2008; Kawasaki A y col., 2008^a; Yang W y col., 2009; Han JW y col., 2009). La combinación de la mayoría de estos estudios en un meta-análisis mostró la solidez de esta asociación (Ji JD y col., 2010). Además de esta señal localizada en el intrón 3, se ha sugerido la existencia de una asociación independiente en el intrón 16, siendo el SNP más asociado el rs3821236 (Abelson AK y col., 2009). Una señal independiente del intrón 3 también se vio en un estudio de casi 10000 pacientes con LES procedentes de múltiples grupos étnicos (afro-americanos, hispano-americanos, asiáticos y europeo-americanos). Dicha señal estaba representada por un haplotipo de riesgo que va del exón 4 al 14 y que está relacionada con la asociación del rs3821236 (Namjou B y col., 2009). Por lo tanto, la contribución de *STAT4* al LES es importante en diferentes grupos étnicos y muestra por lo menos dos señales independientes. Sin embargo, todavía no se ha descrito ningún polimorfismo causal. Sólo se ha observado que los SNPs rs7574865 y rs3821236 se correlacionan con los niveles de expresión de *STAT4*, pero se desconoce cuál es el mecanismo de esta regulación (Abelson AK y col., 2009).

Se ha comunicado la asociación de algunas manifestaciones clínicas de LES con los alelos de riesgo de *STAT4*. En algunos estudios se encontró que el alelo de riesgo de rs7574865 estaba más fuertemente asociado en pacientes con nefritis, anti-ADN de doble cadena, edad temprana de aparición de la enfermedad, desorden hematológico, isquemia cerebrovascular, anticuerpos aPL, y anticuerpos anti-Ro (Taylor KE y col., 2008; Kawasaki A y col., 2008^a; Sigurdsson S y col., 2008^b; Yang W y col., 2009; Svenungsson E y col., 2009; Namjou B y col., 2009). Sin embargo, las asociaciones no fueron concordantes en todos los estudios y el fenotipo asociado está todavía por definir totalmente.

STAT4 codifica para un factor de transcripción que se expresa en linfocitos, macrófagos y DCs, estando implicado en la diferenciación celular, proliferación y apoptosis. *STAT4* es inducido por IFN de tipo I, IL23 y, sobre todo, por IL12. Recientemente, se ha visto que la variante de riesgo de rs7574865 confería una mayor sensibilidad a la señalización por IFN α en células mononucleares de sangre periférica de pacientes con LES (Kariuki SN y col., 2009). *STAT4* se encuentra de forma latente en el citosol, pero tras su activación por citocinas, se fosforila y forma homodímeros que se traslocan al núcleo. En el núcleo, *STAT4* estimula la transcripción de genes específicos incluyendo INF γ , un factor clave en la diferenciación de las células Th1, (Watford WT y col., 2004). Probablemente, *STAT4* sea también necesario para el desarrollo de las células Th17 y la producción por éstas de IL17. En modelos animales, se observó que la deficiencia de *STAT4* favorecía la resistencia a padecer enfermedades autoinmunes (Finnegan A y col., 2002). Además, la inhibición de *STAT4* a través de oligodeoxinucleótidos o oligonucleotidos antisentido es beneficiosa en modelos animales de artritis (Klinman DM y col., 2005; Hildner KM y col., 2007).

3.4.8. *ITGAM*

El gen *ITGAM*, que se encuentra en un *locus* ligado a LES (*I6p11*) (Lee YH y col., 2005; Forabosco P y col., 2006), ha mostrado asociación en estudios específicos y GWAS (SLEGEM y col., 2008; Hom G y col., 2008; Nath SK y col., 2008; Graham RR y col., 2008). El SNP más fuertemente asociado es un SNP no sinónimo del exón 3 de *ITGAM* (rs1143679), que presenta una OR de 1.74 en un meta-análisis de dos cohortes de origen europeo (3818 muestras) y africano (1560 muestras) (Nath SK y col., 2008). El alelo de riesgo codifica para una histidina, en lugar de una arginina, en la posición 77 de la proteína. El papel funcional de este cambio todavía se desconoce, aunque se ha sugerido que podría afectar a la conformación de la proteína y, consecuentemente, alterar la interacción de *ITGAM* con sus ligandos. Otros SNPs de la región mostraron también una fuerte asociación. La mayoría se encontraban en LD con rs1143679 ($r^2 = 0.5-0.85$) (SLEGEM y col., 2008; Hom G y col., 2008), mientras que rs11150610 mostraba una señal independiente (Graham. RR y col., 2008). La asociación de *ITGAM* ha sido fuertemente confirmada en un meta-análisis de nueve cohortes independientes (8211 muestras) de poblaciones europeas e hispanas. Sin embargo, estos polimorfismos son prácticamente monomórficos en asiáticos, constituyendo un claro ejemplo de riesgo de LES específico de población (Han S y col., 2009).

ITGAM codifica para la integrina α -M que dimeriza con la integrina β -2 (ITGB2 o CD18), formando el receptor del complemento 3 (también conocido como CD11b o Mac-1). Este complejo se expresa en la superficie celular de macrófagos, monocitos, DCs y neutrófilos. Las funciones de *ITGAM* son muy variadas e incluyen su participación en la adhesión y trans migración de los leucocitos a través del endotelio (Dunne JL y col., 2003), la activación de

neutrófilos y monocitos (Shappell SB y col., 1990), la apoptosis de neutrófilos (Coxon A y col., 1996) y la eliminación de complejos inmunes mediante fagocitosis. Estas funciones las lleva a cabo a través de su interacción con un amplio rango de ligandos, que incluyen el producto de activación del complemento 3 (C3bi), el fibrinógeno y la molécula 1 de adhesión intracelular (ICAM1). Como los pacientes con LES presentan un defecto en la eliminación de complejos inmunes, se ha sugerido que el polimorfismo de ITGAM podría ser un factor contribuyente. También se ha observado que la unión de C3bi a ITGAM induce la producción del factor de crecimiento transformante β (TGF β) e IL10, que son esenciales para el mantenimiento de la tolerancia, mecanismo igualmente alterado en el LES (Sohn JH y col., 2003).

3.4.9. *C8orf13-BLK*

Uno de los polimorfismos asociados con LES en el estudio GWA de Hom y col., fue un SNP de la región promotora compartida por los genes *BLK* y *C8orf13* (FAM167A) (rs13277113, OR = 1.39) (Hom y col., 2008). Otros 2 GWAS (SLEGEM y col., 2008; Graham RR y col., 2008) encontraron otros 2 SNPs con menor efecto (rs2248932, OR = 1.22; rs2618476, OR = 1.29) en el gen *BLK*. El alelo de riesgo del SNP rs13277113 está asociado con expresión reducida de *BLK* e incrementada de *C8orf13*. No obstante, ni este SNP, ni ningún otro polimorfismo en LD ($r^2 > 0.5$) con él, afectan al sitio de unión de un factor de transcripción ni a otras secuencias nucleicas con función predecible.

C8orf13 es un gen que se expresa de forma ubicua pero que no tiene función conocida. De *BLK* se sabe que codifica una quinasa de tirosina de la familia Src, pero su función todavía no ha sido bien definida. En ratones se ha visto que se expresa en células pro-B tardías y durante su desarrollo a células

plasmáticas (Wasserman R y col., 1995). Por ello, se ha sugerido que estaría implicada en la transición de células pro-B a pre-B (Tretter T y col., 2003). No obstante, ratones deficientes en BLK no muestran defectos en las células B, por lo que su función no está clara (Texido G y col., 2000). En humanos, se ha sugerido que BLK podría actuar en la señalización del receptor de células B (BCR) y de NF κ B, influenciando el desarrollo de células B y los mecanismos de tolerancia y autoinmunidad (Gourh P y col., 2010).

3.4.10. *BANK1*

BANK1 se identificó como un potencial candidato a partir de un estudio GWA de baja densidad (Kozyrev SV y col., 2008). Se identificaron 3 posibles polimorfismos funcionales, que no son independientes. El que mostró una evidencia de asociación más fuerte fue rs17266594 (OR = 0.70), que es un polimorfismo *branch point* antes del exón 2. Este SNP altera el *splicing* de la proteína, dando lugar a una isoforma corta sin exón 2, que podría afectar a su unión con otras proteínas. El segundo polimorfismo asociado (rs10516487, OR = 0.72) es un SNP no sinónimo, localizado a 153 pb del rs17266594 y en LD con él ($r^2 = 0.9$). El tercer SNP de susceptibilidad (rs3733197, OR = 0.81) se localiza en el exón 7, correspondiente al dominio anquirina de la proteína, pero no mostró LD con ninguno de los otros 2 SNPs ($r^2 < 0.4$). La asociación de *BANK1* fue replicada también en un estudio en población china, donde los SNPs rs17266594 y rs3733197 mostraban asociaciones independientes (Chang YK y col., 2009). Se ha sugerido la posible asociación de diversos SNPs de *BANK1* con variables clínicas específicas: daño hematológico, inmunológico y renal y producción de anti-RNP A y anti-Sm (Guo L y col., 2009). Sin embargo, estas asociaciones con la clínica del LES son dudosas, ya que el análisis se hizo comparando los casos positivos para cada criterio con los controles.

BANK1 es una proteína de anclaje que se expresa en células B. Esta proteína media la interacción entre la quinasa de tirosina LYN (*Yamaguchi sarcoma viral oncogene homolog*) y el canal de calcio IP₃R (*inositol 1,4,5-triphosphate receptor*), facilitando la liberación de calcio intracelular, proceso que es importante en la activación de células B (Yokoyama K y col., 2002). Esta función sugiere que las variantes asociadas con LES podrían aumentar el riesgo a sufrir la enfermedad al alterar el umbral de activación de las células B. Sin embargo, BANK1 también parece tener un papel inhibitorio de la respuesta de células B, ya que ratones deficientes en esta proteína muestran una producción elevada de anticuerpos (Aiba Y y col., 2006).

3.4.11. *Iq25.1* y *TNFSF4*

Uno de los SNPs asociados significativamente en el GWAS del SLEGEN fue rs10798269, localizado en el *locus Iq25.1* (SLEGEN y col., 2008), donde no se ha descrito ningún gen. Sólo a unas 133 kb se encuentra un posible gen candidato, *TNFSF4* (también conocido como *OX40L*). Otros tres estudios han mostrado asociación de LES con *TNFSF4*, dos de los cuales pasaban el umbral de significación de los GWAS (Delgado-Vega AM y col., 2009; Gateva V y col., 2009; Graham RR y col., 2009). Resultados similares se obtuvieron en el GWAS realizado con chinos (Chang YK y col., 2009). Este gen se expresa en CPAs y células endoteliales y parece tener un papel en la estimulación de células T CD4⁺. Además, se han observado dos haplotipos en la región *upstream* del gen, uno protector y otro de riesgo, asociados con LES, en un estudio en el que también se encontró que el haplotipo de riesgo correlacionaba con un incremento en la expresión de *TNFSF4* (Cunningham Graham DS y col., 2008^a). Se sugirió que altos niveles de este ligando llevarían a una elevada estimulación de las células T CD4⁺, desestabilizando la tolerancia periférica.

3.4.12. *KIAA1542-IRF7*

El SNP de *KIAA1542* asociado con LES (rs4963128) se ubica en una región donde se ha descrito un polimorfismo de tipo indel (Redon R y col., 2006). El gen, también llamado *PHRF1* (*PHD and ring finger domain-containing protein 1*), se encuentra en un *locus* homólogo al gen codificante de un factor de elongación. Sin embargo, se desconoce la función de la proteína que codifica y se ha sugerido que el efecto de susceptibilidad observado en este *locus* puede deberse a la proximidad del gen *IRF7* (a ~ 20 Kb), que ejerce una función importante en la señalización del IFN de tipo I. De hecho, un SNP de este gen (rs709266) muestra un alto LD ($r^2 = 0.94$) con el rs4963128 (SLEGEN y col., 2008; Gateva V y col., 2009).

3.4.13. *PXK*

PXK codifica para una quinasa de serina-treonina con dominio homólogo a *Phox*, de la que se conocen 5 variantes de *splicing*, 3 de las cuales (*PXK_v1*, *PXK_v2* y *PXK_v4*) se expresan en una amplia variedad de tejidos y una cuarta (*PXK_v3*) que sólo lo hace en leucocitos de sangre periférica. Los dominios homólogos a *Phox* tienen la función de dirigir proteínas hacia la membrana celular, al unirse de forma específica a dominios de fosfoinositol, por lo que son cruciales para diversas funciones celulares (Zou X y col., 2005). El SNP asociado con LES (rs6445975) se localiza en el intrón 5 de este gen, pero se desconoce la posible función del polimorfismo y de la proteína en la susceptibilidad al LES.

3.4.14. *MECP2- IRAK1*

MECP2 e *IRAK1* se encuentran en un mismo bloque haplotípico de *Xq28*. Estudios independientes mostraron de forma convincente asociación con *MECP2* o con *IRAK1*, pero ninguno de ellos ha analizado los dos genes. Por lo tanto, no es posible saber si la asociación se debe a uno u a otro. En el estudio que identificó *MECP2*, el SNP rs17435, en el intrón 2, fue el más fuertemente asociado tanto en población europeo-americana (n = 1080, OR = 1.29) como en coreana (n = 628, OR = 1.58) (Sawalha AH y col., 2008). Otros 3 SNPs (rs1734791, rs1734792, rs1734787) entre los que hay un fuerte LD mostraron también una fuerte asociación. Este bloque define un haplotipo de riesgo y otro de protección. Los mismos haplotipos se encontraron en una colección independiente de 3230 muestras europeo-americanas, donde se confirmó la asociación de *MECP2* con LES (Webb R y col., 2009). El estudio centrado en *IRAK1* se basó en un análisis TDT de 9412 SNPs llevado a cabo en LES infantil, incluyendo 251 familias con la estructura de tríos (Jacob CO y col., 2007). Esta asociación fue confirmada (OR > 1.5) tanto en LES infantil como en adultos, al genotipar 13 SNPs de la región en unas 10000 muestras de 4 grupos étnicos diferentes (europeos, asiáticos, africanos e hispanos) (Jacob CO y col., 2009). Los SNPs más asociados (OR > 1.5) están en un bloque de LD que también define un haplotipo de riesgo y otro de protección. Por tanto, es posible que estos estudios estén describiendo la misma asociación. Además, se sugirió que la dosis de *IRAK1* y *MECP2* podría contribuir a la mayor prevalencia del LES en mujeres, debido a su localización en el cromosoma X (Jacob CO y col., 2009; Sawalha AH y col., 2008).

Tanto *MECP2* como *IRAK1* son buenos genes candidatos. El producto de *MECP2* es una proteína que se une a citosinas metiladas y recluta deacetilasas

de histonas que dan lugar a una conformación de la cromatina inaccesible a la maquinaria transcripcional (Jones PL y col., 1998). Pero se ha sugerido que MECP2 no sólo ejerce una función represora de la transcripción sino que también puede ejercer un efecto opuesto, induciendo la activación de la transcripción, al reclutar al factor de transcripción CREB1 (*cAMP responsive element binding protein 1*) (Chahrour M y col., 2008). Se ha estudiado si el haplotipo de riesgo de *MECP2* afecta a la expresión de sus genes diana. Se observó que estaba correlacionado con expresión diferencial de genes sensibles a metilación en células B, la mayoría sufriendo sobreexpresión y estando algunos de ellos regulados por IFN. Esto último llevó a sugerir que *MECP2* contribuiría a la “firma del IFN” característica del LES (Webb R y col., 2009).

IRAK1 codifica una quinasa de serina-treonina, que está implicada en la transmisión de las señales de receptores de la familia TLR/IL1R. *IRAK1* une varios receptores a la proteína activadora/adaptadora TRAF6 que, una vez activada, regula diversas rutas tanto en la respuesta inmune innata como adquirida. Se ha visto en modelos murinos de LES que *IRAK1* podría tener un papel en la inducción de $\text{INF}\alpha/\gamma$, en la regulación del factor nuclear κB (NF κB) de células T y en la activación de los TLRs.

3.4.15. *TNFAIP3* y región intergénica 6q23

El gen *TNFAIP3* y la región intergénica 6q23 son dos *loci* vecinos que se encuentran en bloques haplotípicos distintos y que están asociados a varias enfermedades autoinmunes incluyendo AR (Wellcome Trust Case Control Consortium, 2007; Thomson W y col., 2007; Dieguez-Gonzalez R y col., 2009) y LES (Graham RR y col., 2008; Musone SL y col., 2008). Sin embargo, los detalles de la asociación no coinciden. El SNP de riesgo que mostraba una

asociación más fuerte con AR (rs6920220) está en la región intergénica *6q23* y sólo mostró una asociación modesta con LES en el estudio de Graham y col., mientras que no estaba asociado en el de Musone y col. De forma similar, el SNP protector para AR en la región intergénica (rs10499194) se vio asociado con LES en el estudio de Musone y col. Por el contrario, las asociaciones en *TNFAIP3* han sido mucho más claras en los estudios de LES que en los de AR. Los dos estudios de LES expusieron la asociación de un haplotipo de riesgo definido por los alelos menores de una serie de SNPs de baja frecuencia altamente correlacionados ($r^2 > 0.98$) que incluyen rs104999197, rs5029939, rs2230926 y rs7749323 (Graham RR y col., 2008; Musone SL y col., 2008). El meta-análisis de ambos estudios y resultados de imputación permitieron definir un haplotipo de riesgo que abarca 109 kb, en el que es difícil determinar si el alelo funcional es el nsSNP rs2230926 u otro de los 14 SNPs que constituyen el haplotipo. Sin embargo, este estudio no consiguió replicar la asociación de la región intergénica *6q23*, debido a los resultados dispares de los dos estudios (Bates JS y col., 2009).

Además, el estudio de Musone y col., encontró asociación con un SNP en la región promotora de un gen implicado en apoptosis, *PERP* (*p53 apoptosis effector related to PMP22*), que se encuentra a unas 200 kb 3' de *TNFAIP3*. Sin embargo, esta asociación no ha sido descrita en ningún estudio adicional.

El bloque de LD de *6q23* abarca unas 60 kb en las que no hay ningún gen conocido, sino sólo el pseudogén *PTPN11*. Este *locus* está flanqueado por los genes del factor de transcripción 3 de oligodendrocitos (*OLIG3*, *oligodendrocyte transcription factor 3*) y *TNFAIP3*. *OLIG3* parece estar implicado en el desarrollo y diferenciación de células neuronales, no siendo un gen candidato para las enfermedades autoinmunes. En contraste, *TNFAIP3*

regula de forma negativa al factor de transcripción NFκB, que sí está implicado en la respuesta inmune, pudiendo tener un papel importante en la enfermedad (Boone DL y col., 2004; Wertz IE y col., 2004).

El gen *TNFAIP3* codifica la proteína A20, un enzima que ubiquitina proteínas de la ruta pro-inflamatoria de NFκB, desencadenada por TLRs, el receptor de IL1 (IL1R), el receptor de TNF (TNFR) y NOD2 (Boone DL y col., 2004; Wertz IE y col., 2004). Las proteínas diana de A20 incluyen las proteínas adaptadoras TRAF6, RIP1, RIP2 e IKKγ/NEMO, que tras ser ubiquitinadas sufren degradación proteosómica (Wertz IE y col., 2004). De este modo, A20 juega un papel importante en la modulación de funciones celulares como la activación celular, la señalización de citocinas y la apoptosis, actuando como molécula anti-inflamatoria (Beyaert R y col., 2000). La importancia de estas funciones de *TNFAIP3* se ha demostrado en ratones deficientes, que desarrollan inflamación multiorgánica severa y caquexia letales (Lee EG y col., 2000). El único polimorfismo asociado que puede tener un papel funcional es rs2230926, que da lugar a un cambio de fenilalanina (alelo mayor) por cisteína (alelo menor) en el residuo 127 de A20. Se observó que este cambio no alteraba la estabilidad de la proteína, pero reducía de forma modesta la inhibición de NFκB, lo que podría llevar a una respuesta inflamatoria exagerada (Musone y col., 2008).

3.4.16. *TNIP1*

TNIP1 mostró asociación significativa con LES en el GWAS de Gateva y col. El SNP asociado, rs7708392, se localiza en un intrón del gen. Otras variantes a unas 21 kb de este gen ya habían mostrado su contribución al riesgo de psoriasis (Nair RP y col., 2009), pero las señales en estas dos enfermedades autoinmunes son independientes ($r^2 = 0.001$). Se sabe que *TNIP1* interacciona

con TNFAIP3 (Heyninck y col., 2003), aunque su papel en la regulación de esta proteína se desconoce. Sin embargo, ambos genes presentan señales de susceptibilidad en distintas enfermedades autoinmunes, por lo que se ha hipotetizado que esta ruta tiene un importante papel en la regulación de la autoinmunidad.

3.4.17. *PRDMI-ATG5*

Dos estudios han observado señales de asociación con LES en el *locus PRDMI-ATG5*. Una de ellas se encuentra en la región intergénica (rs6568431) mientras que la segunda se localiza en *ATG5* (rs2245214, rs573775) (SLEGEN y col., 2008; Gateva V y col., 2009). El bajo LD ($r^2 < 0.1$) y los análisis condicionales sugieren que son señales independientes. *ATG5* está implicado en la autofagia celular, proceso catabólico por el cual el citoplasma y los orgánulos en exceso o deteriorados son secuestrados en lisosomas o vacuolas para su descomposición. Sin embargo, su papel en el LES podría estar relacionado con su participación en la apoptosis dependiente de caspasas, como sugieren los modelos KO del gen. Por otra parte, *PRDM1* también podría tener un papel en el sistema inmune, ya que es un regulador transcripcional de la diferenciación de células T y B (Crotty S y col., 2010).

3.4.18. *JAZF1*

La región promotora del gen *JAZF1* mostró asociación con LES, con $P < 5 \times 10^{-8}$ (Gateva V y col., 2009). La misma variante de susceptibilidad (rs849142) se observó previamente en diabetes tipo 2 (T2D) (Zeggini E y col., 2008). Otro SNP independiente (rs10486567), cercano al gen, también mostró asociación con susceptibilidad al cáncer de próstata (Thomas G y col., 2008),

pero no se asoció a LES (Gateva V y col., 2009). *JAZF1* codifica para un factor de transcripción que regula la expresión de numerosos genes, por lo que se desconoce cuál podría ser su función específica en el LES.

3.4.19. *UHRF1BP1*

Un nsSNP (rs11755393) del gen *UHRF1BP1* es uno de los nuevos *loci* de susceptibilidad al LES identificados en el estudio de Gateva y col. (Gateva y col., 2009). Este cambio aminoacídico se localiza en un posible dominio de unión a UHRF1, un factor de transcripción y metilación ligado a múltiples rutas (Arita K y col., 2008). *UHRF1BP1* se encuentra en una región con alto LD que incluye otros genes, como *SNRPC* (*small nuclear ribonucleoprotein polypeptide C*), que codifica para un polipéptido que forma parte del complejo de procesamiento del ARN. Por tanto, aún está por definir qué gen sería el responsable causal de la asociación en este *locus*.

3.4.20. *IL10*

La IL10 es una importante citocina anti-inflamatoria (Diveu C y col., 2008), por lo que el gen que la codifica es un buen candidato a participar en la susceptibilidad al LES. Los resultados de asociación han sido inconsistentes en los estudios realizados hace unos años. Sin embargo, los últimos estudios que son mucho más sólidos parecen confirmar la asociación con este *locus* (Nath SK y col., 2005; Gateva V y col., 2009). Además, uno de los polimorfismos asociados con LES también contribuye al riesgo de colitis ulcerosa (Franke A y col., 2008) y T1D (Barrett JC y col., 2009), sugiriendo la posibilidad de que el papel de la ruta de IL10 sea similar en estas tres enfermedades.

3.5. Factores genéticos con asociación al LES controvertida

Hay muchos genes que han sido asociados con LES en algún estudio pero que no han sido confirmados con posterioridad. Aquí sólo se tratarán dos que por la solidez de los estudios que los apoyan pueden ser genuinos factores de riesgo aunque no hayan sido todavía replicados de forma definitiva.

3.5.1. *TYK2*

TYK2 es un gen implicado en varias rutas de señalización, entre ellas, la del IFN α . Su relación con la susceptibilidad al LES es controvertida (Sigurdsson S y col., 2005; Cunninghame Graham DS y col., 2007^a; SLEGEN y col., 2008). La asociación de este gen se describió a la vez que la de *IRF5* en un estudio en población escandinava, que analizaba genes de la ruta del IFN (Sigurdsson S y col., 2005). El SNP más asociado fue rs2304256 (OR = 0.63), localizado en el exón 8 del gen. Éste es un nsSNP que codifica para un cambio de valina por fenilalanina en el dominio homólogo *Janus* 4, pero su posible efecto funcional aún no ha sido identificado. En este estudio había un segundo SNP asociado, rs12720356, pero ni éste ni otros 3 SNPs de *TYK2* (rs280500, rs6511696, rs280519) mostraron asociación en ninguno de los GWAS (SLEGEN y col., 2008). Además, la asociación del nsSNP (rs2304256) sólo ha sido confirmada en un pequeño estudio en finlandeses (Hellquist A y col., 2009), mientras que no fue replicada en un estudio TDT con familias inglesas (Cunninghame Graham DS y col., 2007^a). En este último se encontró otro SNP asociado (rs12720270), que también estaba asociado en el estudio finlandés (Hellquist A y col., 2009), pero que no había resultado significativo en las familias escandinavas (Sigurdsson S y col., 2005).

TYK2 es miembro de la familia de quinasas *Janus* y forma parte de la ruta JAK-STAT, desencadenada por los receptores de IFN α , IL6, IL10 o IL12. En las células mononucleares de sangre periférica, TYK2 es activada, junto con JAK1, por el receptor de IFN α (IFNAR), tras unión de su ligando. Esta activación da lugar al reclutamiento y activación de otras señales de transducción, como son las STATs, que tras dimerizar se traslocan al núcleo donde estimulan la expresión de genes de la “firma del IFN” (Kotenko SV y col., 2000; Baechler EC y col., 2003).

3.5.2. *LY9*

LY9 (también conocido como CD229) se encuentra en el intervalo de ligamiento al LES en *1q23* (Cantor RM y col., 2004). Este gen fue el que mostraba asociación más fuerte de entre los estudiados en dicha región en un estudio TDT de 660 familias inglesas y 271 familias canadienses (Cunninghame Graham DS y col., 2008^b). Hasta el momento, ningún otro estudio ha confirmado este factor genético. *LY9* se expresa en la superficie de monocitos, células NK y linfocitos T periféricos. Participa en la adhesión entre células T y CPAs, aunque se conoce poco acerca de su función. El SNP con *P* más significativa (0.002), rs509749, es un nsSNP del exón 8 que codifica para un cambio de valina por metionina en el aminoácido 602 del dominio citoplasmático. Esta región es el sitio de unión de SAP/SH2D1a, una proteína adaptadora implicada en la señalización en células T al reclutar quinasas Src, por lo que podría interferir con esta misma (Cunninghame Graham DS y col., 2008^b).

4. Determinación de los efectos funcionales de los *loci* de susceptibilidad

Como se ha comentado anteriormente, en los últimos dos años se han identificado más de 25 *loci* que predisponen al LES, reflejando el éxito que han tenido los GWAS en esta enfermedad (Moser KL y col., 2009). Estos resultados han permitido una mayor comprensión de las alteraciones inmunológicas responsables del LES, ya que muchos de los genes asociados tienen papeles importantes en el sistema inmune. Sin embargo, otros no tienen una función conocida o son *loci* localizados en regiones del ADN sin ninguna función identificable (**Figura 7**). Además, incluso en los genes con función conocida no se sabe todavía cuál es el polimorfismo causal y cuáles son sus consecuencias específicas que motivan la susceptibilidad al LES. Esta ignorancia es una de las dificultades en la aplicabilidad de los descubrimientos genéticos.

Para tratar de avanzar en estos aspectos es necesario descubrir los polimorfismos causales e identificar sus efectos. Éstos pueden ser desde cambios en la estructura, función o estabilidad de las proteínas a alteraciones en las secuencias reguladoras, ya sean sitios de unión de factores de transcripción o de miRNAs, o sitios afectados por cambios epigenéticos. Un ejemplo de un posible polimorfismo causal y de su posible efecto es el de PD1.3 en el gen *PDCDI*, que fue considerado muy convincente aunque en la actualidad está en duda.

	Células	Rutas	Genes
Respuesta Inmune Innata	Células dendríticas Macrófagos	Señalización TLR/IFN	IRF5 STAT4 SPP1 IRAK1 TREK1
		Señalización TNF/NFκB	TNFAIP3
Función/ Activación Linfocitaria	Células T autorreactivas	Señalización Células T	HLA-DR PTPN22 STAT4 PDCD1 IRAK1 TNFSF4
	Células B autorreactivas	Señalización Células B	HLA-DR BLK BANK1 FCGR2B LYN
Eliminación complejos inmunes	Macrófagos Neutrófilos	Fagocitosis	FCGR3A FCGR3B CRP ITGAM
		Complemento	C4A C4B C2 C1q
Otros		Apoptosis	ATG5 STAT4
		Ubiquitinación	UBE2L3 TNFAIP3
		Metilación ADN	MECP2
		Adhesión celular	ITGAM
		Desconocida	PXK ICA1 SCUBE1 NMNAT2 XKR6 KIAA1542

Figura 7. Genes asociados con LES y su función en el sistema inmune (Moser KL y col., 2009).

PD1.3 fue el SNP más fuertemente asociado en un gran estudio de LES anterior a los GWAS. Éste sería un ejemplo de SNP que puede afectar al sitio de unión de un factor de transcripción. La sustitución del alelo G por el alelo A de PD1.3 afecta al cuarto nucleótido del sitio de unión consenso (TGT/cGGT) para el factor de transcripción RUNX1 (*runt-related transcription factor 1*) (**Figura 8**). Este SNP se localiza en el cuarto intrón de *PDCD1* en una secuencia que muestra cuatro repeticiones en *tandem* de 40 nucleótidos cada una. Se propuso que estas repeticiones podrían actuar como *enhancer* y, por tanto, tener un papel importante en la regulación transcripcional de *PDCD1* (Prokunina L y col., 2002). Los *enhancer* son elementos reguladores que potencian la transcripción y que son capaces de actuar a grandes distancias del gen e independientemente de su posición y orientación con respecto al gen. La hipótesis de que el intrón 4 de *PDCD1* funcione como *enhancer* se basa en la predicción informática de

múltiples sitios de unión a diferentes factores de transcripción implicados en la diferenciación hematopoyética e inflamación: RUNX1, NFκB1 (o p50) y factores de unión al E-box (**Figura 8**).

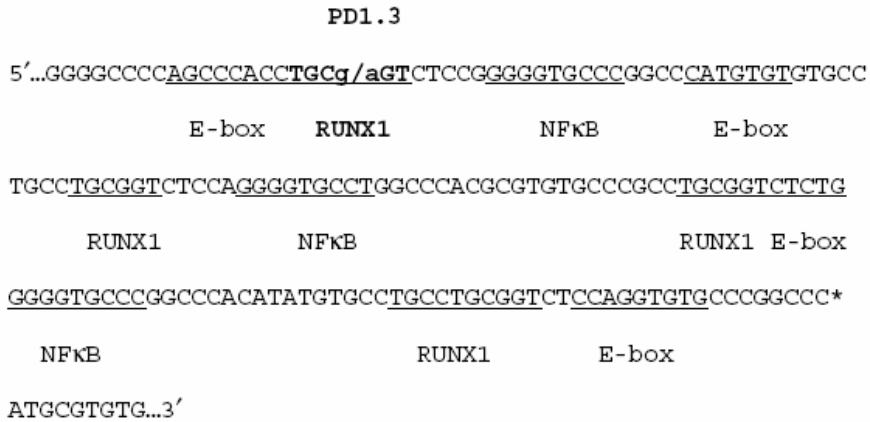


Figura 8. Parte de la secuencia del *enhancer* hipotético en el intrón 4 de *PDCDI* (Prokunina L y col., 2002)

Sólo se ha publicado un experimento que apoyaba el papel funcional de PD1.3. Se trata de un ensayo de cambio de movilidad electroforética (EMSA) en el que se observó unión específica de RUNX1 a un oligonucleótido sintético con el alelo G de PD1.3, y desaparición de la unión al oligonucleótido cuando llevaba el alelo A (Prokunina L y col., 2002). Por tanto, se sugirió que el SNP tendría un efecto regulador de *PDCDI*, probablemente disminuyendo su expresión. En el mismo estudio también se observaron diferencias en la expresión del ARN mensajero de *PDCDI* en células mononucleares de sangre periférica activadas entre los pacientes AA y GA y los GG. Sin embargo, estos resultados no fueron concluyentes debido a grandes variaciones interindividuales (Prokunina L y col., 2002). Por tanto, el papel funcional de

este SNP todavía no está claro y su posible papel en la regulación de la expresión de *PDCD1* requiere ser confirmada.

El factor de transcripción RUNX1, que se uniría a PD1.3, pertenece a una familia de factores de transcripción nucleares, CBF (*core binding factor*), que realizan funciones importantes en el desarrollo normal y en enfermedades. Concretamente, RUNX1 regula la expresión de genes implicados en la hematopoyesis y en la diferenciación de células mieloides y linfoides. La secuencia consenso de unión de RUNX1 fue descrita a partir de la selección de oligonucleótidos sintéticos que presentaban afinidad por el factor de transcripción (Meyers S y col., 1993). Esta secuencia, TGT/cGGT, pertenece a un grupo de secuencias identificadas previamente como elementos críticos para la actividad de los *enhancers* del poliomavirus, del virus SV40 (*simian virus 40*) y del retrovirus murino SL3-3 (Thornell A y col., 1988; Kamachi Y y col., 1990). Pero además, es funcional en múltiples *enhancers* y promotores celulares de genes implicados en la hematopoyesis y en la función del sistema inmune, como en *enhancers* del *TCR* (Hernandez-Munain C y col., 1995; Bruhn L y col., 1997) y en los promotores de los genes *IL3* (Uchida H y col., 1997), factor estimulador de colonias de granulocitos-macrófagos (*GM-CSF*) (Frank R y col., 1995), *CD3ε* (Hallberg B y col., 1992), adaptador monocítico (*Mona/Gads*) (Guyot B y col., 2003), granzima B (*GRZB*) (Wargnier A y col., 1995), proteína 1α inflamatoria de macrófagos (*MIP1α*) (Bristow CA y col., 2003), *IL2* e *INFγ* (Ono M y col., 2007). Así, el papel de RUNX1 en todas estas regiones reguladoras apoya la hipótesis de que la secuencia donde se encuentra PD1.3 pueda actuar como un *enhancer*, pero todavía no ha sido comprobado.

5. El problema de la heredabilidad no explicada

Los GWAS han sido una potente herramienta para la investigación del componente genético de las enfermedades complejas, identificando gran cantidad de variantes genéticas de susceptibilidad. La mayoría de asociaciones encontradas hasta ahora muestran ORs ≤ 1.5 (**Figura 6**). Este pequeño efecto y el gran número de genes de susceptibilidad que se están encontrando indica que en el LES participan múltiples factores genéticos ya que cada uno de ellos confiere un pequeño aumento en el riesgo a padecer la enfermedad. Sin embargo, a pesar de ser numerosos explican sólo una parte pequeña de la heredabilidad (proporción del fenotipo atribuible a factores genéticos). Así, en el estudio GWA del SLEGEN se estimó que *PXK*, *HLA*, *IRF5*, *KIAA1542* e *ITGAM* explican tan sólo un 15 % de la heredabilidad (SLEGEN y col., 2008). Estos resultados llevan a preguntarse dónde puede residir la heredabilidad todavía no identificada.

Varias han sido las sugerencias sobre las distintas razones que podrían explicar la falta de detección de una proporción considerable de la heredabilidad. En primer lugar, es posible que aún no se hayan identificado muchas de las variantes comunes de pequeño efecto. Otra posibilidad es que las variantes causales por descubrir sean variantes de baja frecuencia (1-5 %) o variantes raras (frecuencia < 0.5 %), que no están bien cubiertas por los GWAS actuales. Además, es posible que estas variantes poco frecuentes tengan un efecto más notable. Sin embargo, su efecto no puede ser muy fuerte ya que no han sido detectadas por los estudios de ligamiento (Manolio TA y col., 2009). Un ejemplo de variantes de este tipo son las variantes raras en *TREX1* (Lee-Kirsch MA y col., 2007) que ya han sido comentadas. Otra posible fuente de heredabilidad no explicada son las variantes estructurales, como las CNVs.

Hasta hace poco, estas variaciones no se evaluaban de forma explícita ni estaban bien cubiertas en los GWAS. Sólo recientemente se han desarrollado varias aproximaciones para integrar el análisis de las CNVs en los GWAS, lo que ha llevado al descubrimiento de su efecto en desórdenes neuropsiquiátricos, enfermedad de Crohn y obesidad. En el LES todavía no se han utilizado estas nuevas herramientas pero estudios previos ya habían detectado fuertes asociación con dos CNVs. Una en el factor del complemento *C4* (Manderson AP y col., 2004) y otra en *FCGR3B* (Willcocks LC y col., 2008). Por último, otros posibles contribuyentes a la heredabilidad no explicada, que los GWAS tienen poca potencia para detectar, son las interacciones gen-ambiente o gen-gen, fenómenos conocidos como epistasis (Manolio TA y col., 2009) y que se desarrollará a continuación.

La epistasis se puede analizar desde dos perspectivas diferentes, biológica y estadística. En la perspectiva biológica, la epistasis es el resultado de la interacción entre moléculas que participan en una misma ruta, de modo que el efecto de un gen en un fenotipo depende de otro u otros genes o de algún factor ambiental. En contraste, la epistasis estadística se define como una desviación de la aditividad en un modelo matemático que relaciona múltiples genotipos y/o variables ambientales con un fenotipo. Por lo tanto, no existe una relación directa entre estas dos formas de considerar la epistasis. Es posible que la demostración de epistasis por análisis estadísticos no se correlacione con una relación biológica clara, y también que algunas epistasis biológicas no den lugar a epistasis desde el punto de vista estadístico. Dada la herencia poligénica del LES y el pequeño efecto de los polimorfismos asociados, el desafío está en demostrar si existen interacciones que pudieran explicar el riesgo de LES más allá de la acumulación de los riesgos individuales definidos por cada uno de los polimorfismos de susceptibilidad.

Como parte del estudio GWA llevado a cabo por el consorcio SLEGEN, se exploró si había evidencias de epistasis entre los genes asociados, pero no las encontraron (SLEGEN y col., 2008). Otros estudios analizaron la interacción entre genes que forman parte de la misma ruta funcional. Así, *STAT4* e *IRF5* son miembros de la ruta de IFN de tipo I, por lo que sus polimorfismos de susceptibilidad serían candidatos a interaccionar. Lo que se observó en un estudio de 500 pacientes y 500 controles suecos fue que con cada alelo de riesgo del rs7582694 de *STAT4* (que define el haplotipo de riesgo y tiene $r^2 = 1$ con el rs7574865) y de dos SNPs (indel del promotor y rs10488631) de *IRF5* la OR se incrementaba en 1.82 (Sigurdsson S y col., 2008^b). Un efecto incrementado al combinar estos dos factores genéticos se describió en otro estudio de LES con muestras europeas e hispanas (Abelson AK y col., 2009), así como en pacientes con síndrome de Sjögren (SS) (Nordmark G y col., 2009) y en pacientes con esclerosis sistémica cutánea difusa (Dieudé P y col., 2009). Sin embargo, los aumentos de riesgo observados en estos estudios no han detectado desviación de la suma de susceptibilidad que confiere cada alelo por sí mismo, por lo que no hay evidencias de interacción epistática entre estos dos factores.

Otro factor que participa en la ruta del IFN es *TYK2*, por lo que también se ha hipotetizado su posible interacción con *IRF5*. De hecho, un estudio encontró epistasis significativa ($P = 0.014$) entre los SNPs rs10954213 de *IRF5* y rs2304256 de *TYK2*. La combinación genotípica que más contribuyó a la interacción era la de los homocigotos para los alelos de riesgo de ambos SNPs (OR = 2.73) (Hellquist A y col., 2009). Aún así, el tamaño muestral de este estudio no es muy grande (300 pacientes y 300 controles), por lo que sería necesaria la confirmación de este resultado por estudios independientes de mayor tamaño.

Por último, en un estudio aún sin publicar, se ha examinado también la interacción entre moléculas que participan en la señalización de las células B. Concretamente, se buscaron interacciones genéticas entre *BANK1* y ~100000 SNPs de un GWAS. *BANK1* mostró interacción estadísticamente significativa con 29 genes, entre los que se incluían *BLK* y el receptor 2 de inositol 1,4,5-trifosfato (*ITPR2*). Además, se confirmó la interacción física entre *BANK1* y *BLK* y que el alelo de riesgo del SNP rs17266594 de *BANK1*, que da lugar a una variante de *splicing* sin dominio de unión, reduce su interacción con *BLK* (Castillejo-Lopez C y col., 2009).

Por tanto, las evidencias de epistasis entre factores de susceptibilidad al LES son pocas por el momento. La de *TYK2* e *IRF5* es débil, posiblemente por el pequeño tamaño muestral empleado, mientras que la de *BANK1*, *BLK* e *ITPR2* parece más prometedora ya que ha sido confirmada en dos cohortes de gran tamaño. Además, la interacción de *BANK1* y *BLK* se ha explicado biológicamente al observar como estas dos proteínas también interaccionan físicamente y que la interacción se relaciona con una variante de *splicing* (Castillejo-Lopez C y col., 2009).

Un campo inexplorado es el de las epistasis entre polimorfismos que no están asociados con LES así como el de las interacciones con el ambiente. Por todo ello, todavía es posible que existan muchos más ejemplos de epistasis y que tengan una contribución significativa a la heredabilidad del LES.

6. La hipótesis de los factores genéticos comunes a múltiples enfermedades autoinmunes a la luz de los GWAS

Las enfermedades autoinmunes se caracterizan porque se inician por la pérdida de tolerancia a autoantígenos. Algunos de los mecanismos implicados en los procesos patogénicos son específicos de cada una de ellas, pero también comparten otros. Entre estos últimos destacan los mecanismos de desregulación de la activación de linfocitos T y B y las rutas que llevan a la inflamación. Un ejemplo clásico de los efectos de estos mecanismos compartidos es el de los ratones NOD, un modelo de T1D, que además presentan susceptibilidad a diferentes enfermedades autoinmunes o a desarrollar anticuerpos característicos de dichas enfermedades (Johansson AC y col., 2003). Así mismo, en humanos se ha observado que algunas enfermedades autoinmunes concurren en un mismo individuo o en una misma familia más frecuentemente que lo esperado (Baranzini SE y col., 2009). Esto ha llevado a sugerir que, al igual que comparten mecanismos patogénicos, podrían compartir parte de su etiología, en concreto los factores genéticos que predisponen a la autoinmunidad.

Los estudios de ligamiento han mostrado regiones que influyen la susceptibilidad a varias enfermedades autoinmunes. El ejemplo más claro es el del *HLA*, que es el factor que mayor influencia tiene en la susceptibilidad a autoinmunidad. Además, hay otros genes que han mostrado asociación convincente con diferentes desórdenes autoinmunes. Tres ejemplos notables son el antígeno 4 de linfocitos T citotóxicos (*CTLA4*), *PTPN22* y *STAT4*. Estos genes están implicados en mecanismos inmunogenéticos comunes y, por lo tanto, tienen un efecto pleiotrópico en los fenotipos autoinmunes (Serrano NC y col., 2006). En los últimos años se han llevado a cabo más de 30 GWAS en más de 7 enfermedades autoinmunes, que han identificado genes asociados con más

de una enfermedad. Entre los nuevos asociados se pueden destacar el receptor α de IL2 (*IL2RA*), asociado con esclerosis múltiple y T1D, *IL12B*, asociado con psoriasis y enfermedad de Crohn y *TNFAIP3*, asociado con AR y LES (Baranzini SE y col., 2009).

En el ámbito de las enfermedades reumáticas, existen algunas evidencias de que el LES y la AR pueden compartir parte del componente genético. Así, el riesgo a padecer AR de los hermanos de pacientes con LES estuvo aumentado unas 3-5 veces con respecto al riesgo de la población general en un estudio (Alarcón-Segovia D y col., 2005). También estas dos enfermedades se solapan en algunos pacientes raros, dándosele a esta condición el nombre de “*rhus*” (Panush RS y col., 1988).

La AR es una enfermedad compleja en la que el componente genético juega un papel destacable (Aho K y col., 1986; Silman AJ y col., 1993). El factor genético más fuerte está en el *HLA* (Seldin MF y col., 1999). Concretamente, esta asociación se debe a la tercera región hipervariable de la cadena DR β 1, especialmente los aminoácidos 70-74, que forman parte del lugar de unión al péptido (Nepom GT y col., 1989). El segundo factor genético más importante en población caucásica es el polimorfismo R620W de *PTPN22*, que aumenta el riesgo en un 40-80 % (Hinks A y col., 2006). Otros *loci* de susceptibilidad destacados son la región *6q23*, *STAT4*, *TRAF1/C5*, *IL2RB*, *PRKCQ*, el miembro 5A de la familia de las quinesinas (*KIF5A*), el miembro 3 de la familia AF4/FMR2 (*AFF3*), *CD40*, *CTLA4*, *IL2-IL21* y *MMEL1* (*membrane metallo-endopeptidase-like 1*) (Barton A y col., 2009). Muchos de estos genes son importantes en la regulación del sistema inmune, estando implicados en la regulación de las células T y de las rutas de señalización o en la apoptosis (Barton A y col., 2008). Por lo tanto, el *HLA*, *PTPN22*, *6q23* y

STAT4, son factores genéticos compartidos por la AR y el LES (Zhernakova A y col., 2009; Barton A y col., 2009) y existe la posibilidad de que no sean los únicos. A continuación se aborda con más detalle la asociación con AR de uno de estos factores que comparten ambas enfermedades.

6.1. *TNFAIP3* y región intergénica 6q23

La región intergénica 6q23 se observó asociada con AR en dos GWAS llevados a cabo en población inglesa y europeo-americana (Wellcome Trust Case Control Consortium, 2007; Plenge RM y col., 2007^b). Esta asociación se debía a dos señales independientes, una de riesgo (rs6920220) y la otra de protección (rs10499194, rs13207033). Este nuevo *locus* de susceptibilidad a AR se confirmó en cohortes de muestras independientes (Thomson W y col., 2007; Plenge RM y col., 2007^a; Dieguez-Gonzalez R y col., 2009; Orozco G y col., 2009^b). Es de resaltar que la señal protectora solamente ha sido observada en pacientes positivos para anticuerpos frente a péptidos citrulinados (ACPA). Además, también se han visto evidencias de asociación con AR en el gen vecino *TNFAIP3*, siendo estas señales estadísticamente independientes a las de la región 6q23 (Dieguez-Gonzalez R y col., 2009; Orozco G y col., 2009^b). Estos resultados sumados a la falta de LD entre la región intergénica y *TNFAIP3* proporcionan evidencias que soportan el efecto independiente de estas dos regiones.

Estudios en soriasis (Nair RP y col., 2009), enfermedad de Crohn (Wellcome Trust Case Control Consortium, 2007), T1D (Fung EY y col., 2009) y LES (Musone SL y col., 2008; Graham RR y col., 2008) sugieren que *TNFAIP3* también juega un papel en estas enfermedades autoinmunes y apoyan la hipótesis de la existencia de una base genética compartida. Sin embargo,

algunas de estas señales no están bien recogidas en nuestro estudio previo. Las más interesantes son las asociaciones encontradas en LES debido a su fuerte efecto y a que una se corresponde con un nsSNP que podría tener un papel causal en la susceptibilidad a la enfermedad.

En resumen, en los últimos años se han publicado una gran cantidad de estudios de asociación de gran tamaño que han permitido la identificación de un gran número de factores de susceptibilidad a LES. Estos nuevos *loci* necesitan ser confirmados por estudios independientes. Además, para determinar la importancia de estos factores en el LES es preciso identificar los posibles polimorfismos causales y determinar su papel funcional. Así, se ha sugerido a PD1.3 como posible polimorfismo causal en el gen *PDCDI*, aunque este papel sólo había sido explorado mediante un EMSA. Por tanto, es necesario realizar otros estudios funcionales que puedan confirmar esta hipótesis. Por otra parte, el componente genético no sólo influencia la susceptibilidad al LES, sino que también puede influenciar la variabilidad fenotípica de la enfermedad. La exploración de este efecto permitirá una mayor comprensión de la patogenia del LES. Por último, los estudios de asociación de LES y AR mostraron que una parte de los factores de susceptibilidad están asociados con ambas enfermedades autoinmunes. Así pues, los nuevos *loci* de susceptibilidad al LES deberían considerarse como candidatos en los estudios de asociación de AR.

HIPÓTESIS

Hipótesis

1. El LES es una enfermedad compleja con un importante componente genético. Los estudios de asociación caso-control son una herramienta útil para la identificación de dichos factores genéticos pero los resultados que obtienen necesitan ser confirmados en colecciones de muestras independientes y por distintos laboratorios.
2. Uno de los factores de susceptibilidad a LES se encuentra en el gen *PDCDI*. En este gen se propone que el SNP de susceptibilidad, PD1.3, tiene un papel funcional en la expresión de *PDCDI* debido a su localización en una región con estructura de *enhancer* y a que altera el sitio de unión al factor de transcripción RUNX1.
3. Algunas de las diferencias en las manifestaciones clínicas de los pacientes con LES pueden explicarse por influencia del componente genético.
4. El LES comparte mecanismos patogénicos con otras enfermedades autoinmunes. Por lo tanto se propone que parte de sus factores genéticos serán también factores de susceptibilidad a la AR.

OBJETIVOS

Objetivos

1. Confirmar hallazgos significativos de asociación con susceptibilidad a LES en una gran cohorte independiente de muestras europeas, mediante un estudio de asociación caso-control.
2. Demostrar mediante estudios funcionales que PD1.3 es un polimorfismo causal en *PDCDI*.
3. Explorar la influencia del componente genético en la variabilidad clínica del LES, desde dos puntos de vista: el de los factores de susceptibilidad al LES y el del fondo genético diferenciado de las dos grandes subpoblaciones de europeos, la del norte y la del sur.
4. Comprobar si los nuevos factores de susceptibilidad al LES se asocian con susceptibilidad a la AR.

PROCEDIMIENTO EXPERIMENTAL Y RESULTADOS

Procedimiento experimental y resultados

Esquema general

1. Factores genéticos de LES

Publicación 1:

- Suarez-Gestal M, Calaza M, Endreffy E, Pullmann R, Ordi-Ros J, Sebastiani GD, Ruzickova S, Santos MJ, Papasteriades C, Marchini M, Skopouli FN, Suarez A, Blanco FJ, D'Alfonso S, Bijl M, Carreira P, Witte T, Migliaresi S, Gomez-Reino JJ, Gonzalez A and the European Consortium of SLE DNA Collections. *Replication of recently identified systemic lupus erythematosus genetic associations: a case-control study*. Arthritis Res Ther 2009, 11: R69.

Publicación 2:

- Suarez-Gestal M, Calaza M, Gonzalez A. *Lack of interaction between systemic lupus erythematosus-associated polymorphisms in TYK2 and IRF5*. J Rheumatol. 2010; 37: 676-677.

2. Estudio funcional de PD1.3

Publicación 3:

- Suarez-Gestal M, Ferreiros-Vidal I, Ortiz JA, Gomez-Reino JJ, Gonzalez A. *Analysis of the functional relevance of a putative regulatory SNP of PDCD1, PD1.3, associated with systemic lupus erythematosus*. Genes Immun. 2008; 9: 309-315.

3. Influencia del componente genético en la variabilidad clínica del LES

Publicación 4:

- Suarez-Gestal M, Calaza M, Liz M, Ordi-Ros J, Balada E, Bijl M, Kallenberg CG, Papasteriades C, Kappou-Rigatou I, Carreira P, Skopouli FN, Mavromati M, Schmidt RE, Witte T, Endreffy E, Kovacs A, Marchini M, Scorza R, Migliaresi S, Sebastiani GD, Santos MJ, Vinagre F, Suarez A, Gutierrez C, Rego I, Blanco FJ, Barizzzone N, D'Alfonso S, Pullmann Jr R, Pullmann R, Ruzickova S, Dostal C, Gomez-Reino JJ and Gonzalez A. *Weak association of systemic lupus erythematosus clinical features with susceptibility alleles: a case-only study*.

Publicación 5:

- Suarez-Gestal M, Calaza M, Witte T, Papasteriades C, Marchini M, Migliaresi S, Kovacs A, Ordi-Ros J, Bijl M, Santos MJ, Ruzickova S, Pullmann R, Carreira P, Skopouli FN, D'Alfonso S, Sebastiani GD, Suarez A, Blanco FJ, Gomez-Reino JJ, Gonzalez A and the European Consortium

of SLE DNA Collections. *Differences of systemic lupus erythematosus clinical features between two large European population subgroups.*

4. Componente genético común a LES y AR

Publicación 6:

- Suarez-Gestal M, Calaza M, Dieguez-Gonzalez R, Perez-Pampin E, Pablos JL, Navarro F, Narváez J, Marengo JL, Herrero-Beaumont G, Fernandez-Gutierrez B, Lamas JR, Rodriguez de la Serna A, Ortiz AM, Carreño L, Cañete JD, Caliz R, Blanco FJ, Balsa A, Gomez-Reino JJ and Gonzalez A. *Rheumatoid arthritis does not share most of the newly identified systemic lupus erythematosus genetic factors.* Arthritis Rheum. 2009. 60 (9): 2558-2564

Publicación 7:

- Suarez-Gestal M, Calaza M, Herrero-Beaumont G, Cañete JD, Ortiz AM, Marengo JL, Caliz R, Fernandez-Gutierrez B, Narváez J, Rodriguez de la Serna A, Dieguez-Gonzalez R, Carreño L, Balsa A, Blanco FJ, Perez-Pampin E, Pablos JL, Navarro F, Gomez-Reino JJ and Gonzalez A. *Follow-up case-control study of the TNFAIP3 association with rheumatoid arthritis: possible role of antinuclear antibody status and differences with other diseases.*

1. Factores genéticos de LES

Research article

Open Access

Replication of recently identified systemic lupus erythematosus genetic associations: a case-control study

Marian Suarez-Gestal¹, Manuel Calaza¹, Emöke Endreffy², Rudolf Pullmann³, Josep Ordi-Ros⁴, Gian Domenico Sebastiani⁵, Sarka Ruzickova⁶, Maria Jose Santos^{7,8}, Chryssa Papasteriades⁹, Maurizio Marchini¹⁰, Fotini N Skopouli¹¹, Ana Suarez¹², Francisco J Blanco¹³, Sandra D'Alfonso¹⁴, Marc Bijl¹⁵, Patricia Carreira¹⁶, Torsten Witte¹⁷, Sergio Migliaresi¹⁸, Juan J Gomez-Reino^{1,19}, Antonio Gonzalez¹ for the European Consortium of SLE DNA Collections

¹Laboratorio de Investigacion 10 and Rheumatology Unit, Hospital Clinico Universitario de Santiago, Santiago de Compostela 15706, Spain

²Paediatrics Department, Albert Szent-Györgyi Medical and Pharmaceutical Centre, University of Szeged, Szeged 6721, Hungary

³Institute of Clinical Biochemistry, Martin Faculty Hospital, Jessenius Medical Faculty, Kollárova 2, 036 59 Martin, Slovakia

⁴Internal Medicine, Research Laboratory in Autoimmune Diseases, Hospital Vall d'Hebron, 08035 Barcelona, Spain

⁵Ospedale S Camillo-Forlanini, U O Complessa di Reumatologia, 00151 Roma, Italy

⁶Molecular Biology and Immunogenetics Department, Institute of Rheumatology, 128 50 Prague 2, Czech Republic

⁷Rheumatology Department, Hospital Garcia de Orta, Almada, Portugal

⁸Rheumatology Research Unit, Instituto Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Portugal

⁹Department of Histocompatibility and Immunology, Evangelismos Hospital, 10676 Athens, Greece

¹⁰Clinical Immunology, University of Milan and Fondazione IRCCS Ospedale Maggiore Policlinico, Mangiagalli e Regina Elena, 20122 Milan, Italy

¹¹Pathophysiology Department, Athens University Medical School, Athens 115 27, Greece

¹²Department of Functional Biology, Hospital Universitario Central de Asturias, Universidad de Oviedo, Oviedo 33006, Spain

¹³INIBIC-CH Universitario A Coruña, 15006 A Coruña, Spain

¹⁴Dept Medical Sciences and IRCAD, Eastern Piedmont University, 28100 Novara, Italy

¹⁵Department of Rheumatology and Clinical Immunology, University Medical Center Groningen, 9713 Groningen, The Netherlands

¹⁶Rheumatology Unit Hospital 12 de Octubre, 28041 Madrid, Spain

¹⁷Division of Clinical Immunology, Department of Internal Medicine of the Hannover Medical School, D-30625 Hannover, Germany

¹⁸Rheumatology Unit, Second University of Naples, 81100 Naples, Italy

¹⁹Department of Medicine, University of Santiago de Compostela, Santiago de Compostela, 15706, Spain

Corresponding author: Antonio Gonzalez, anlugon@hotmail.com

Received: 24 Mar 2009 Revisions requested: 29 Apr 2009 Revisions received: 8 May 2009 Accepted: 14 May 2009 Published: 14 May 2009

Arthritis Research & Therapy 2009, **11**:R69 (doi:10.1186/ar2698)

This article is online at: <http://arthritis-research.com/content/11/3/R69>

© 2009 Suarez-Gestal et al.; licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Introduction We aimed to replicate association of newly identified systemic lupus erythematosus (SLE) loci.

Methods We selected the most associated SNP in 10 SLE loci. These 10 SNPs were analysed in 1,579 patients with SLE and 1,726 controls of European origin by single-base extension. Comparison of allele frequencies between cases and controls was done with the Mantel-Haenszel approach to account for heterogeneity between sample collections.

Results A previously controversial association with a SNP in the *TYK2* gene was replicated (odds ratio (OR) = 0.79, $P = 2.5 \times$

10^{-5}), as well as association with the X chromosome *MECP2* gene (OR = 1.26, $P = 0.00085$ in women), which had only been reported in a single study, and association with four other loci, *1q25.1* (OR = 0.81, $P = 0.0001$), *PXK* (OR = 1.19, $P = 0.0038$), *BANK1* (OR = 0.83, $P = 0.006$) and *KIAA1542* (OR = 0.84, $P = 0.001$), which have been identified in a genome-wide association study, but not found in any other study. All these replications showed the same disease-associated allele as originally reported. No association was found with the *LY9* SNP, which had been reported in a single study.

Conclusions Our results confirm nine SLE loci. For six of them, *TYK2*, *MECP2*, *1q25.1*, *PXK*, *BANK1* and *KIAA1542*, this

BANK1: B-cell scaffold protein with ankyrin repeats 1; **BLK**: B-lymphoid tyrosine kinase; **GWA**: genome-wide association; **IL**: interleukin; **ITGAM**: integrin alpha M; **LY9**: lymphocyte antigen 9; **MECP2**: methyl CpG binding protein 2; **OR**: odds ratio; **PXK**: PX domain containing serine/threonine kinase; **SLE**: systemic lupus erythematosus; **SLEGEN**: International Consortium for Systemic Lupus Erythematosus Genetics; **SNP**: single nucleotide polymorphism; **STAT4**: signal transducer and activator of transcription; **TYK2**: tyrosine kinase 2.

replication is important. The other three loci, *ITGAM*, *STAT4* and *C8orf13-BLK*, were already clearly confirmed. Our results also suggest that *MECP2* association has no influence in the sex

bias of SLE, contrary to what has been proposed. In addition, none of the other associations seems important in this respect.

Introduction

Systemic lupus erythematosus (SLE) is a complex autoimmune disease of wide variability in its manifestations and clinical evolution that characteristically involves multiple autoantibodies against ubiquitous nuclear antigens. Its genetic component is very significant, as shown by a sibling recurrence risk ratio of 20 and a 10-fold excess in SLE concordance between monozygotic twins over dizygotic twins [1,2].

Linkage studies have indicated that this genetic component is due to multiple low-penetrance common genetic factors [1]. Only a few factors had consistently been demonstrated until 2008: the class II HLA alleles, low-affinity receptors for the constant fraction of IgG, and the *PTPN22* and *IRF5* genes. This scenario has been dramatically improved by new technologies and genome resources [2]. Four genome-wide association (GWA) studies were published in 2008 [3-6] that, together with other large-scale studies, have greatly enlarged the number of convincing SLE-associated loci. Not all of the newly described findings, however, have attained the same degree of confirmation [2]. Some of them are already definitively confirmed by replication in different sample collections by the same authors and also by independent authors in separate studies (Table 1). In this group are the SLE associations with the *ITGAM* [3,4,6,7], *STAT4* [3,4,6,8-12] and *C8orf13-BLK* regions [3,4,6]. Other findings are very solid but they still require confirmation by independent studies. In this group are the associated loci that were only reported in a single GWA study but not in the other studies, such as *BANK1* [5], *PXK* [3], *KIAA1542* [3] and *1q25.1* [3], or those that were reported in a single large study but not in any of the four GWA studies, such as *MECP2* [13] and *LY9* [14]. Finally, the *TYK2* association is more controversial because it was found in a large study with Scandinavian families [15], partially replicated in a large study of UK families [16], and excluded in one of the GWA studies [3].

In the present paper, therefore, we have analysed SLE association to each of these loci in more than 1,500 SLE patients and 1,700 controls – and all of them except *LY9* have been clearly replicated. In addition, we have found that many of these loci are also important for SLE in men where data from previous reports is almost completely absent.

Materials and methods

Sample collection

We used DNA samples from SLE patients and ethnically matched healthy controls of 16 collections from nine Euro-

pean countries (see Table S1 in Additional data file 1). Most of these samples have already been described [17]. Two new sample collections were from Asturias, Spain and Almada, Portugal. Each recruiting centre was asked for about 100 SLE patients and 100 ethnically matched controls. A total of 1,579 cases and 1,726 controls were obtained in this way. All SLE patients met the revised American College of Rheumatology classification criteria [18]. Clinical characteristics of the patients are provided in Table S2 in Additional data file 1. Patients and controls gave written informed consent. Sample collection was approved by the respective ethical committees.

Genotyping

We selected a SNP for each of the 10 associated loci that we intended to replicate (Table 1). The SNPs were selected because they were strongly associated with SLE or because they were described as probable causal polymorphisms. These 10 SNPs were amplified in a single PCR with the Qiagen Multiplex PCR kit (Qiagen, Chatsworth, CA, USA) with 20 ng genomic DNA and 0.2 μ M of each primer (for primers and probes, see Table S3 in Additional data file 1). The PCR products were purified by digestion with Exonuclease I (Epicentre, Madison, WI, USA) and shrimp alkaline phosphatase (GE Healthcare, Barcelona, Spain). Purified PCR products were genotyped by single-base extension with the SNaPshot Multiplex Kit (Applied Biosystems, Foster City, CA, USA) and specific probes. After a second purification with shrimp alkaline phosphatase (GE Healthcare), samples were analysed in the Abi Prism 3130xl Genetic Analyzer (Applied Biosystems) and genotypes assigned by the GeneMapper software. All genotype calls were manually reviewed and conflicting results were liberally re-assayed or re-genotyped by sequencing with the Big Dye Ready Reaction Kit v 3.1 (Applied Biosystems). Sequence reactions followed the kit manufacturer protocol and were also analysed in the Abi Prism 3130xl Genetic Analyzer.

Statistical analysis

Some of the sample collections in our study have already been used for the analysis of specific associations included in this project. They have been excluded from the relevant analyses to avoid data duplication; this circumstance is detailed in Table S4 in Additional data file 1, where raw genotype data from each sample collection are reported. Hardy-Weinberg equilibrium tests in control samples were performed with Haploview with a threshold of 0.05 uncorrected for multiple tests [19]. Other statistical analyses were carried out in a customized version of the Statistica 7.0 program (StatSoft, Tulsa, OK, USA).

Table 1

Newly systemic lupus erythematosus-associated loci that were examined with previous evidence of association

Locus	Chr.	Location	SNP	Alleles ^a	OR (95% CI)	Associated	Sample size	Population	Reference
<i>ITGAM</i>	16	Exon 3	rs1143679	G/A	1.78 (1.6 to 2.0)	Yes	3,818	European American	[7]
					1.55 (1.2 to 2.0)	Yes	1,289	African American	[7]
					2.07 (1.3 to 3.4)	Yes	271	Gullah	[7]
					Other SNPs		Yes	[3,4,6]	
<i>STAT4</i>	2	Intron 3	rs7574865	G/T	1.5 (1.2 to 1.8)	Yes	3,057	European	[3]
					1.50 (1.4 to 1.7)	Yes	4,651	European American	[4]
					1.55 (1.3 to 1.8)	Yes	2,287	European American	[8]
					1.61 (1.4 to 1.9)	Yes	2,495	Japanese	[9]
					1.62 (1.2 to 2.2)	Yes	565	Colombians	[10]
					1.56 (1.4 to 1.7)	Yes	3,958	European	[12]
Other SNPs		Yes	[6,11]						
<i>C8orf13-BLK</i>	8	Intergenic	rs13277113	G/A	1.39 (1.3 to 1.5)	Yes	6,301	European American	[4]
						Yes			[3,6]
<i>TYK2</i>	19	Exon 8	rs2304256	C/A	0.625 (0.5 to 0.8)	Yes	966	Swedish/Finnish	[15]
					105:127 ^b	No	380 ^c	European/Indo-Pakistani	[16]
					Other SNPs		No	[3]	
<i>MECP2</i>	X	Intron 2	rs17435	A/T	1.58 (1.3 to 1.9)	Yes	1,364	Korean	[13]
					1.29 (1.1 to 1.5)	Yes	2,160	European	[13]
<i>1q25.1</i>	1	Intergenic	rs10798269	G/A	0.82 (0.8 to 0.9)	Yes	6,728	European	[3]
<i>BANK1</i>	4	Intron 1	rs17266594	T/C	0.74 (0.6 to 0.9)	Yes	927	Scandinavian	[5]
					0.70 (0.5 to 0.9)	Yes	576	German	[5]
					0.63 (0.5 to 0.8)	Yes	450	Italian	[5]
					0.78 (0.6 to 0.9)	Yes	1,136	Spanish	[5]
					0.58 (0.4 to 0.8)	Yes	620	Argentinian	[5]
<i>KIAA1542</i>	11	Intron 4	rs4963128	G/A	0.78 (0.7 to 0.9)	Yes	6,728	European	[3]
<i>PXK</i>	3	Intron 5	rs6445975	T/G	1.25 (1.2 to 1.4)	Yes	6,728	European	[3]
<i>LY9</i>	1	Exon 8	rs509749	A/G	377:403 ^d	Yes	510 ^c	UK Caucasian	[14]
					237:251 ^d	No	270 ^c	Canadian	[14]

Loci ordered as presented in Table 2. The SNPs selected for replication are detailed. Chr, chromosome; CI, confidence interval; OR, odds ratio.

^aMajor/minor alleles. ^bTransmitted/untransmitted. ^cNumber of families. ^dObserved:expected.

Comparison of cases and controls was carried out with the Mantel–Haenszel approach because allele frequency differences are probable between sample collections even if specific effects on the phenotype are constant. Spurious false positive or false negative results therefore become likely if the allele differences are not accounted for. To avoid this, the Mantel–Haenszel approach combines effect sizes taken as the odds ratio (OR) in each stratum allowing for heterogeneity in allele frequencies. This approach provides an accurate combined statistic if the heterogeneity of effect sizes, evaluated

with the Breslow–Day test, is excluded. Significant heterogeneity of effects is therefore excluded by the Breslow–Day test and allele frequency heterogeneity is accounted for with the Mantel–Haenszel approach. These analyses were also conducted after stratifying the samples by gender. Univariate logistic regression models were used to test the fit to the data of additive, recessive and dominant genetic models. Statistical power was estimated with the Power and sample size calculations software [20].

Results

A total of 1,579 SLE patients and 1,726 controls from 16 European collections were available for study (Tables S1 and S2 in Additional data file 1). The genotyping call rate was 99.9% and the genotypes in controls were in Hardy-Weinberg equilibrium. Individual collection data for each SNP is shown in Tables S4 and S5 in Additional data file 1. Combined analysis of the SNP effects across our sample collections was performed with the Mantel-Haenszel approach, which is a method correcting for variability in allele frequencies between collections provided that the effect sizes (that is, ORs) are not significantly divergent. This condition was fulfilled because no significant heterogeneity in OR was detected for any of the SNPs (Table 2, final column).

The combined data showed significant differences between SLE cases and controls for eight of the nine SNPs located in autosomal chromosomes (Table 2). All of the significant differences between cases and controls were in the same direction as originally reported (Tables 1 and 2). We found association of the four SNPs that have been reported in a single GWA and not yet replicated by independent studies: rs10798269 in *1q25.1* (OR = 0.81, $P = 0.00013$), rs6445975 in *PXK* (OR = 1.19, $P = 0.0038$), rs17266594 in *BANK1* (OR = 0.83, $P = 0.0062$) and rs4963128 in *KIAA1542* (OR = 0.84, $P = 0.0011$). There was also significant association of two of the three SNPs that were described in large studies but that were not observed in any of the GWA studies: rs2304256 in *TYK2* (OR = 0.79, $P = 2.5 \times 10^{-5}$) and rs17435 in *MECP2* (analysis of this SNP was performed separately in women and men because this gene is in chromosome X; see below). Only rs509749 in *LY9* was similar in cases and controls. Our study had sufficient power (80%) to detect association at this SNP

with an effect size equivalent to OR > 1.15 with $P < 0.05$ (or OR > 1.23 for $P < 0.001$).

In addition to these important results for replication, we found association with the three loci that have already been replicated in GWA studies: rs1143679 in *ITGAM* (OR = 1.70, $P = 1.1 \times 10^{-16}$), rs7574865 in *STAT4* (OR = 1.62, $P = 2.4 \times 10^{-12}$) and rs13277113 in *C8orf13-BLK* (OR = 1.34, $P = 5.1 \times 10^{-7}$). The effect sizes of these three association signals (that is, their ORs) were larger than for all the other signals, perhaps explaining the more consistent replication of their association. Genotype comparisons for the different SNPs were concordant with an additive genetic model and yielded very similar results to the allele frequency analyses (data not shown).

Combined analysis was also conducted in women (Table 3). This was particularly necessary for the *MECP2* SNP rs17435, located in the X chromosome. This SNP showed a significant difference between SLE women and control women and with the same disease-associated allele as previously reported (OR = 1.26, $P = 0.00085$). The SNPs placed in the autosomes showed similar results to those obtained in the unstratified analysis. There were only less significant P values due to the smaller sample size, but the effect sizes (expressed as ORs) remained largely unchanged. The *BANK1* SNP was not associated in women, but this was the SNP with fewer available samples because we have excluded from this analysis the sample collections that have previously been reported (power was 0.68 for $P = 0.05$ and OR = 0.78, which was previously reported in Spanish samples) [5].

No previous detailed information of men with SLE has been published for any of these associated loci, although in a report

Table 2

Combined analysis of allele frequency differences between SLE cases and controls for nine autosomal loci

SNP (locus)	Minor allele frequency (%) ^a		Mantel – Haenszel analysis		Breslow – Day test
	SLE cases	Controls	OR (95% CI)	P value	P value
rs1143679 (<i>ITGAM</i>)	23.2 (730/3,152)	15.1 (521/3,448)	1.70 (1.5 to 1.9)	1.1×10^{-16}	0.5
rs7574865 (<i>STAT4</i>)	32.8 (709/2,162)	23.2 (485/2,092)	1.62 (1.4 to 1.9)	2.4×10^{-12}	1.0
rs13277113 (<i>C8orf13-BLK</i>)	30.9 (874/2,824)	25.7 (776/3,024)	1.34 (1.2 to 1.5)	5.1×10^{-7}	1.0
rs2304256 (<i>TYK2</i>)	23.3 (733/3,152)	27.8 (960/3,450)	0.79 (0.7 to 0.9)	2.5×10^{-5}	0.6
rs10798269 (<i>1q25.1</i>)	27.3 (861/3,158)	31.8 (1,098/3,452)	0.81 (0.7 to 0.9)	0.00013	0.2
rs17266594 (<i>BANK1</i>)	24 (526/2,192)	27.6 (624/2,260)	0.83 (0.7 to 0.9)	0.0062	0.4
rs4963128 (<i>KIAA1542</i>)	30.3 (955/3,150)	34.0 (1,173/3,448)	0.84 (0.8 to 0.9)	0.0011	0.2
rs6445975 (<i>PXK</i>)	27.3 (772/2,824)	24.2 (734/3,034)	1.19 (1.1 to 1.3)	0.0038	0.7
rs509749 (<i>LY9</i>)	43.1 (1,359/3,154)	43.8 (1,513/3,452)	0.97 (0.9 to 1.1)	0.5	0.4

Loci ordered by decreasing effect size (odds ratio (OR)). All results refer to the minor allele of each SNP, which is indicated in Table 1. CI, confidence interval; SLE, systemic lupus erythematosus. ^aData presented as percentage (number of minor alleles/total number of alleles).

Table 3**Combined analysis of allele frequency differences between SLE women and control women**

SNP (locus)	Minor allele frequency (%) ^a		Mantel – Haenszel analysis		Breslow – Day test
	SLE cases	Controls	OR (95% CI)	P value	P value
rs1143679 (<i>ITGAM</i>)	22.3 (621/2,782)	15.1 (329/2,182)	1.67 (1.4 to 2.0)	2.0×10^{-11}	0.6
rs7574865 (<i>STAT4</i>)	33.1 (636/1,920)	24.0 (317/1,322)	1.60 (1.4 to 1.9)	8.4×10^{-9}	0.8
rs13277113 (<i>C8orf13-BLK</i>)	30.9 (777/2,514)	25.7 (475/1,848)	1.33 (1.2 to 1.5)	5.4×10^{-5}	0.9
rs2304256 (<i>TYK2</i>)	22.9 (638/2,782)	27.2 (592/2,180)	0.81 (0.7 to 0.9)	0.0022	0.2
rs17435 (<i>MECP2</i>)	26.7 (744/2,784)	22.8 (498/2,186)	1.26 (1.1 to 1.4)	0.00085	0.7
rs10798269 (<i>1q25.1</i>)	27.1 (754/2,786)	32.3 (705/2,182)	0.77 (0.7 to 0.9)	6.2×10^{-5}	0.4
rs17266594 (<i>BANK1</i>)	23.9 (464/1,938)	26.9 (386/1,436)	0.87 (0.7 to 1.0)	0.077	0.4
rs4963128 (<i>KIAA1542</i>)	30.4 (845/2,778)	33.9 (741/2,182)	0.85 (0.8 to 1.0)	0.011	0.2
rs6445975 (<i>PXK</i>)	26.7 (670/2,514)	23.5 (436/1,854)	1.22 (1.1 to 1.4)	0.0067	0.9
rs509749 (<i>LY9</i>)	43.3 (1,206/2,784)	44.0 (962/2,184)	0.98 (0.9 to 1.1)	0.7	0.5

Loci ordered as presented in Table 2. All results refer to the minor allele of each SNP, which is indicated in Table 1. CI, confidence interval; OR, odds ratio; SLE, systemic lupus erythematosus. ^aData presented as percentage (number of minor alleles/total number of alleles).

describing association of the *ITGAM* gene it was indicated that results were not different between women and men [7]. This lack of information is probably due to the rarity of men suffering from SLE. In our analysis, we have considered all male data together without stratifying for sample collection due to the low number of men in each collection (Table 4). Results in men were similar to results in women, with the possible exception of the rs1143679 in *ITGAM* (OR = 2.08 versus 1.67; $P =$

0.03). Some SNPs were not associated in men (in the *TYK2*, *1q25.1*, *BANK1* and *LY9* loci), but statistical power of this subgroup analysis was low, ranging from 0.19 for rs17266594 in *BANK1* to 0.25 for rs2304256 in *TYK2* among the nonassociated SNPs (power was estimated for $P = 0.05$ and OR = 1.2).

Table 4**Comparison of SNP allele frequencies between SLE men and control men**

SNP (locus)	Minor allele frequency (%) ^a		Mantel – Haenszel analysis	
	SLE cases	Controls	OR (95% CI)	P value
rs1143679 (<i>ITGAM</i>)	27.9 (83/298)	15.7 (181/1,154)	2.08 (1.5 to 2.8)	1.2×10^{-6}
rs7574865 (<i>STAT4</i>)	31.7 (64/202)	21.5 (159/740)	1.69 (1.2 to 2.4)	0.0025
rs13277113 (<i>C8orf13-BLK</i>)	31.0 (88/284)	25.1 (267/1,064)	1.34 (1.0 to 1.8)	0.045
rs2304256 (<i>TYK2</i>)	26.8 (80/298)	28.8 (334/1,158)	0.91 (0.7 to 1.2)	0.5
rs17435 (<i>MECP2</i>) ^b	29.2 (42/144)	18.5 (105/568)	1.82 (1.2 to 2.8)	0.0046
rs10798269 (<i>1q25.1</i>)	28.0 (84/300)	30.7 (355/1,158)	0.88 (0.7 to 1.2)	0.4
rs17266594 (<i>BANK1</i>)	26.2 (56/214)	28.4 (227/798)	0.89 (0.6 to 1.3)	0.5
rs4963128 (<i>KIAA1542</i>)	28.0 (84/300)	34.9 (403/1,156)	0.73 (0.6 to 1.0)	0.025
rs6445975 (<i>PXK</i>)	32.1 (86/268)	24.8 (265/1,068)	1.43 (1.1 to 1.9)	0.016
rs509749 (<i>LY9</i>)	40.6 (121/298)	42.9 (496/1,156)	0.91 (0.7 to 1.2)	0.5

Loci ordered as presented in Table 2. All results refer to the minor allele of each SNP, which is indicated in Table 1. No stratified analysis by sample collection was done due to the small number of patients with systemic lupus erythematosus (SLE) in each collection. CI, confidence interval; OR, odds ratio. ^aData presented as percentage (number of minor alleles/total number of alleles). ^bThese results are of carrier analysis because *MECP2* is in the X chromosome and there is a single allele in each man.

Discussion

Our aim has been to contribute to the definition of consistent SLE genetic factors derived from recent sound studies: four of the associations have been described in a GWA study, but not in a second GWA study or any other study; another two associations were identified in large studies, but not in any of the GWA studies; and one association is more controversial (Table 1). Our results are highly reassuring because all of the associations, except one from the group not found in any GWA study, were replicated with clarity and showed the same disease-associated allele as originally reported. This high degree of reproducibility is a fundamental change that large studies have brought to genetic research of SLE and other complex diseases [2,21]. This change allows a bright future for the investigation of the genetic component of SLE.

The most remarkable result from the present study has probably been the association signal observed with the rs2304256 nonsynonymous SNP of *TYK2* (OR = 0.79) because this has been a controversial SLE genetic factor. The rs2304256 SNP introduces a valine to phenylalanine change in the Janus homology domain 4 of *TYK2* whose functional relevance has not yet been tested. This nonsynonymous SNP showed the strongest association among the 11 *TYK2* SNPs studied in Scandinavian families [15], but was not associated in a study of UK families [16]. This latter study, however, found association with another *TYK2* SNP (rs12720270) that was not associated in the Scandinavian study. Finally, the International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEN) GWA study excluded association with the rs12720270 SNP (the rs2304256 SNP was not included in the GWA panels) [3]. Our results are important in this context because they show a significant association that confirms the role of the rs2304256 nonsynonymous SNP. In addition, combined analysis of all available data show a clear SLE association ($P = 2.10 \times 10^{-11}$) that is stronger than the required for genome-wide significance.

TYK2 is a Janus-family tyrosine kinase that is bound to cytokine receptors and becomes activated after ligand binding. Deficiency of *TYK2* leads to defects of multiple cytokine pathways, including type I interferon, IL-6, IL-10, IL-12, and IL-23, and to impaired T-helper type 1 differentiation and accelerated T-helper type 2 differentiation [22]. Only future research will indicate which of these pathways is critically affected by the *TYK2* risk allele.

Following in importance is the association of *MECP2* because our results provide replication and indicate that a previous assumption about the role of this genetic factor in contributing to the sex bias in SLE is questionable. Sawalha and colleagues considered the X-chromosome methyl CpG binding protein 2 coding gene (*MECP2*) as a possible SLE genetic factor based on two features: SLE predominance in women and abnormal regulation of methylation-sensitive T-cell genes

in SLE [13]. *MECP2* could be involved in both phenomena because this gene is in the X chromosome and participates in DNA methylation. Sawalha and colleagues found association with several SNPs in women from two ethnic groups, Korean and European (OR for rs17435 = 1.58 and 1.29, respectively) [13]. The association we have found in women (OR = 1.26) is very similar to that reported in their European sample, providing strong confirmatory evidence. This replication is important for the status of *MECP2* due to the lack of association signals in the SLE GWA studies.

In addition, we have found that the *MECP2* SNP is also associated with SLE in men (OR = 1.82, $P = 0.0046$), which was not previously known. This result seriously undermines the hypothesized role of *MECP2* in SLE gender bias. In retrospect, lack of sex specificity is congruent with experiments that showed *MECP2* is not expressed in the inactivated X chromosome of women [23], which implies expression levels in men and women should be equivalent. Future research should aim to establish whether any of the SLE-associated SNPs in *MECP2* has a functional effect and to find evidence of the hypothesized relationship between altered methylation of T-cell genes in SLE and *MECP2*. In addition, it is even unclear whether the causal polymorphism affects *MECP2* because SLE association has also been reported with genetic variants in a neighbour gene, *IRAK1*, which is a key mediator in the signalling pathways of Toll-like receptors/IL-1R [24].

The rs10798269 SNP in the *1q25.1* locus, the rs4963128 SNP in the *KIAA1542* gene and the rs6445975 SNP in the *PXK* gene were reported in the SLEGEN GWA study [3] with P values below 2×10^{-7} , but they were not reported in Hom and colleagues' GWA study [4] and none of them has yet been replicated in any other study. The three SNPs were associated with SLE in our study, with effect sizes that are similar to those reported (OR = 0.81 versus 0.82 for the *1q25.1* SNP, 0.84 versus 0.78 for the *KIAA1542* SNP, and 1.19 versus 1.25 for the *PXK* SNP). None of these three SNPs has any predictable functional effect. In addition, the rs10798269 SNP in the *1q25.1* locus is far from any known transcript and the *PXK* and *KIAA1542* genes are of unknown function. The *KIAA1542* gene, however, is about 20 kb away from the *IRF7* gene and in linkage disequilibrium with it, raising the possibility that this association could be related with *IRF7* function [3]. Our replication of these associations increases the need for research aimed to the identification of their functional effects.

We have also found a significant association with the rs17266594 in the *BANK1* gene. This SLE genetic factor has been identified in a low-resolution GWA study in a Swedish sample and replicated in other European sample collections in the same study [5], but it was not found in any of the high-resolution GWA studies and has not yet been replicated by other groups. Our results provide this independent replication, although with a more modest effect (OR = 0.83 in our study

versus 0.70 in Kozyrev and colleagues [5]). The causal polymorphism can be the rs17266594 SNP itself, which seems to alter splicing efficiency of *BANK1*, or two *BANK1* nonsynonymous SNPs of possible damaging effect. Linkage disequilibrium between these three SNPs has prevented dissection of their relationship to SLE susceptibility [5]. *BANK1* codes for a B-cell scaffold protein with ankyrin repeats that is implicated in B-cell receptor-mediated signalling.

The rs509749 SNP of *LY9* is the only SNP that was not replicated in our study. We selected this SNP because it seems to explain the 1q23 SLE-linked locus according to a large family-based study [14]. 1q23 is one of the most consistently described SLE loci in linkage studies (and its syntenic region in the mouse lupus models) [1]. Examination of SNPs all along this locus showed stronger association with the rs509749 SNP [14]. This SNP has a predictable impact in protein function and is associated with changes in the proportion of specific T-cell subsets [14]. All this evidence made the rs509749 SNP a good candidate for replication in our view, even if the level of significance of the SLE association was notably lower than the reported for the other nine SNPs studied here ($P = 0.002$). Lack of replication of this SNP in contrast with replication of the other nine SNPs provides support for the direct relationship between very low P values obtained in sound studies and the reproducibility of genetic association findings [21].

The most associated SNPs in our samples were the three that were already confirmed previous to our study. These three SNPs were associated with SLE in at least three large studies. The largest effect was observed with a nonsynonymous SNP in the third exon of the *ITGAM* gene (rs1143679, OR = 1.70) [3,4,6,7]. This nonsynonymous SNP was the most associated in one of the previous studies (with very similar effect, OR = 1.74) [7], and has been hypothesized to disturb ITGAM interaction with its ligands, but still no functional evidence is available. Another clearly established association [3,4,6,8-12] was the second strongest in our study: SNP rs7574865 in the third intron of the *STAT4* gene (OR = 1.62). This association seems stronger in patients with a severe phenotype [12]; however, no functional polymorphism has been identified in this locus. The next strongest association (OR = 1.34) was with the rs13277113 SNP, which has been reported in the GWA study of Hom and colleagues [4], with a similar effect (OR = 1.39). This SNP is located between *C8orf13* (of unknown function) and *BLK* (B-lymphoid tyrosine kinase), two genes that are transcribed in opposite directions. No functional variant has been identified in this locus, but the risk allele of this SNP correlates with low mRNA levels of *BLK* and high levels of *C8orf13*, raising the possibility that either of these two effects could be related with SLE. Graham and colleagues found association with a strongly linked SNP in the *BLK* gene [6], while the SLEGEN GWA study found association with an unlinked SNP in this locus, suggesting the possibility of two independent genetic factors [3].

In addition, we have found that most examined SLE-associated SNPs seem to be shared between women and men. Results are not definitive given the small number of men in the patient group. This lack of differential association is important because we do not know definitively the causes of the female preference of SLE. Lack of detailed gender analysis in previous genetic reports is regrettable because only aggregation of data from multiple studies will allow us to know whether genetic factors contribute to this sex bias.

Conclusions

In summary, our study has provided independent replication of nine SLE-associated loci, six of them of confirmatory importance because they have not yet been independently replicated by other groups (1q25.1, *MECP2*, *KIAA1542*, *PXK* and *BANK1*) or because their association was controversial (*TYK2*). These results bring the number of strongly confirmed associated loci to 13. Replication in independent studies is indispensable for considering a genetic factor in this category, although the common use of multiple case-control sets inside the same study or of large sample collections has increased the chances of replication [2]. Some other promising associations have been discovered [6,25], or await sufficient independent replication [2], but it is already certain that the genetic component of SLE is especially rich in genetic factors with effects above the detectable level with current studies (OR = 1.15 to 1.25). We are therefore now in a phase of exciting discoveries in this field. There still remain formidable challenges, however, because it is necessary to transform the information we obtain into useful knowledge and, as has been discussed above, we have very few clues regarding the meaning of the identified SLE associations. Future studies should try to identify the causal variants and to determine their effect at molecular, cellular and disease levels, including the assessment of their role in the different SLE phenotypes and the probable similar effect in women and men.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MS-G participated in design of the study, in genotyping the samples, in interpretation of the results and in writing the manuscript. MC participated in the statistical analysis and in the interpretation of results. EE, RP, JO-R, GDS, SR, MJS, CP, MM, FNS, AS, FJB, SD'A, MB, PC, TW and SM participated in the acquisition of clinical data and collection of samples and in the analysis and interpretation of results. JJG-R coordinated the acquisition of clinical data and collection of samples and participated in the analysis and interpretation of results. AG participated in the design of the study and in the coordination of acquisition of clinical data and collection of samples, and supervised genotyping, statistical analysis, interpretation of results and writing of the manuscript. All authors read and approved the final manuscript.

Authors' information

Other contributors to the European Consortium of SLE DNA Collections: Attila Kovacs (Albert Szent-Györgyi Medical and Pharmaceutical Centre, University of Szeged, Hungary); Rudolf Pullmann Jr (Gerontology Research Center, National Institute on Aging, Baltimore, MD, USA); Eva Balada (Hospital Vall d'Hebron, Barcelona, Spain); Ctibor Dostal (Institute of Rheumatology, Prague, Czech Republic); Filipe Vinagre (Hospital Garcia de Orta, Almada, Portugal and Instituto Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Portugal); Iris Kappou-Rigatou (Evangelismos Hospital, Athens, Greece); Raffaella Scorza (University of Milan and Fondazione IRCCS Ospedale Maggiore Policlinico, Mangiagalli e Regina Elena, Milan, Italy); Maria Mavromati (Athens University Medical School, Athens, Greece); Carmen Gutierrez (Hospital Universitario Central de Asturias, Universidad de Oviedo, Spain); Ignacio Rego (INIBIC-CH Universitario A Coruña, Spain); Nadia Barizzone (Eastern Piedmont University, Novara, Italy); Cees G Kallenberg (University Medical Center Groningen, The Netherlands); and Reinhold E Schmidt (Hannover Medical School, Hannover, Germany).

Note

A report published in *Arthritis & Rheumatism* after publication of this manuscript provided further confirmation of the association of *MECP2* with SLE [26].

Additional files

The following Additional files are available online:

Additional file 1

A Word file containing Table S1 that lists the origin and female percentage of the DNA sample collections, Table S2 that lists the clinical characteristics of the patients with SLE, Table S3 that lists the primers and probes used for genotyping the 10 SNPs, Table S4 that lists the genotype counts for each of the 10 SNPs detailed for each of the sample collections, and Table S5 that lists the minor allele percentages for each of the 10 SNPs for each of the sample collections.

See <http://www.biomedcentral.com/content/supplementary/ar2698-S1.doc>

Acknowledgements

The authors thank Carmen Pena-Pena for providing outstanding technical assistance. MS-G is the recipient of a FPU predoctoral bursary of the Spanish Ministry of Education. The present work was supported by Fondo de Investigación Sanitaria of the Instituto de Salud Carlos III (Spain), grants 04/1651 and 06/0620 that are partially financed by the Fondo Europeo de Desarrollo Regional program of the European Union, by grants from the Xunta de Galicia, and by BMBF KN Rheuma grant C2.12 (to TW).

References

- Kelly JA, Moser KL, Harley JB: **The genetics of systemic lupus erythematosus: putting the pieces together.** *Genes Immun* 2002, **3**(Suppl 1):S71-S85.
- Rhodes B, Vyse TJ: **The genetics of SLE: an update in the light of genome-wide association studies.** *Rheumatology (Oxford)* 2008, **47**:1603-1611.
- International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEn), Harley JB, Alarcón-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, Moser KL, Tsao BP, Vyse TJ, Langefeld CD, Nath SK, Guthridge JM, Cobb BL, Mirel DB, Marion MC, Williams AH, Divers J, Wang W, Frank SG, Namjou B, Gabriel SB, Lee AT, Gregersen PK, Behrens TW, Taylor KE, Fernando M, Zidovetzki R, Gaffney PM, Edberg JC, Rioux JD, et al.: **Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PTK, KIAA1542 and other loci.** *Nat Genet* 2008, **40**:204-210.
- Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, Lee AT, Chung SA, Ferreira RC, Pant FV, Ballinger DG, Kosoy R, Demirci FY, Kamboh MI, Kao AH, Tian C, Gunnarsson I, Bengtsson AA, Rantapää-Dahlqvist S, Petri M, Manzi S, Seldin MF, Rönblom L, Syvänen AC, Criswell LA, Gregersen PK, Behrens TW: **Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX.** *N Engl J Med* 2008, **358**:900-909.
- Kozryev SV, Abelson AK, Wojcik J, Zaghloul A, Linga Reddy MV, Sanchez E, Gunnarsson I, Svenungsson E, Sturfelt G, Jönsen A, Truedsson L, Pons-Estel BA, Witte T, D'Alfonso S, Barizzone N, Danieli MG, Gutierrez C, Suarez A, Junker P, Lastrup H, González-Escribano MF, Martin J, Abderrahim H, Alarcón-Riquelme ME: **Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus.** *Nat Genet* 2008, **40**:211-216.
- Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM, Burt NP, Guiducci C, Parkin M, Gates C, Plenge RM, Behrens TW, Wither JE, Rioux JD, Fortin PR, Graham DC, Wong AK, Vyse TJ, Daly MJ, Altshuler D, Moser KL, Gaffney PM: **Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus.** *Nat Genet* 2008, **40**:1059-1061.
- Nath SK, Han S, Kim-Howard K, Kelly JA, Viswanathan P, Gilkeson GS, Chen W, Zhu C, McEver RP, Kimberly RP, Alarcón-Riquelme ME, Vyse TJ, Li QZ, Wakeland EK, Merrill JT, James JA, Kaufman KM, Guthridge JM, Harley JB: **A nonsynonymous functional variant in integrin- α (M) (encoded by ITGAM) is associated with systemic lupus erythematosus.** *Nat Genet* 2008, **40**:152-154.
- Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, de Bakker PI, de JM, Lee HS, Batliwalla F, Li W, Masters SL, Booty MG, Carulli JP, Padyukov L, Alfreðsson L, Klareskog L, Chen W, Amos CI, Criswell LA, Seldin MF, Kastner DL, Gregersen PK: **STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus.** *N Engl J Med* 2007, **357**:977-986.
- Kobayashi S, Ikar K, Kaneko H, Kochi Y, Yamamoto K, Shimane K, Nakamura Y, Toyama Y, Mochizuki T, Tsukahara S, Kawaguchi Y, Terai C, Hara M, Tomatsu T, Yamanaka H, Horiuchi T, Tao K, Yasutomo K, Hamada D, Yasui N, Inoue H, Itakura M, Okamoto H, Kamatani N, Momohara S: **Association of STAT4 with susceptibility to rheumatoid arthritis and systemic lupus erythematosus in the Japanese population.** *Arthritis Rheum* 2008, **58**:1940-1946.
- Palomino-Morales RJ, Rojas-Villarraga A, González CI, Ramírez G, Anaya JM, Martín J: **STAT4 but not TRAF1/C5 variants influence the risk of developing rheumatoid arthritis and systemic lupus erythematosus in Colombians.** *Genes Immun* 2008, **9**:379-382.
- Sigurdsson S, Nordmark G, Garnier S, Grundberg E, Kwan T, Nilsson O, Eloranta ML, Gunnarsson I, Svenungsson E, Sturfelt G, Bengtsson AA, Jönsen A, Truedsson L, Rantapää-Dahlqvist S, Eriksson C, Alm G, Göring HH, Pastinen T, Syvänen AC, Rönblom L: **A risk haplotype of STAT4 for systemic lupus erythematosus is over-expressed, correlates with anti-dsDNA and shows additive effects with two risk alleles of IRF5.** *Hum Mol Genet* 2008, **17**:2868-2876.
- Taylor KE, Remmers EF, Lee AT, Ortmann WA, Plenge RM, Tian C, Chung SA, Nititham J, Hom G, Kao AH, Demirci FY, Kamboh MI, Petri M, Manzi S, Kastner DL, Seldin MF, Gregersen PK, Behrens TW, Criswell LA: **Specificity of the STAT4 genetic association**

- for severe disease manifestations of systemic lupus erythematosus. *PLoS Genet* 2008, **4**:e1000084.
13. Sawalha AH, Webb R, Han S, Kelly JA, Kaufman KM, Kimberly RP, Alarcón-Riquelme ME, James JA, Vyse TJ, Gilkeson GS, Choi CB, Scofield RH, Bae SC, Nath SK, Harley JB: **Common variants within MECP2 confer risk of systemic lupus erythematosus.** *PLoS ONE* 2008, **3**:e1727.
 14. Cunninghame Graham DS, Vyse TJ, Fortin PR, Montpetit A, Cai YC, Lim S, McKenzie T, Farwell L, Rhodes B, Chad L, Hudson TJ, Sharpe A, Terhorst C, Greenwood CM, Wither J, Rioux JD, CaNIOS GenES Investigators: **Association of LY9 in UK and Canadian SLE families.** *Genes Immun* 2008, **9**:93-102.
 15. Sigurdsson S, Nordmark G, Göring HH, Lindroos K, Wiman AC, Sturfelt G, Jönsen A, Rantapää-Dahlqvist S, Möller B, Kere J, Koskenmies S, Widén E, Eloranta ML, Julkunen H, Kristjansdóttir H, Steinsson K, Alm G, Rönnblom L, Syvänen AC: **Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus.** *Am J Hum Genet* 2005, **76**:528-537.
 16. Cunninghame Graham DS, Akil M, Vyse TJ: **Association of polymorphisms across the tyrosine kinase gene, TYK2 in UK SLE families.** *Rheumatology (Oxford)* 2007, **46**:927-930.
 17. Ferreira-Neira I, Calaza M, Alonso-Perez E, Marchini M, Scorza R, Sebastiani GD, Blanco FJ, Rego I, Pullmann R Jr, Pullmann R, Kallenberg CG, Bijl M, Skopouli FN, Mavromati M, Migliaresi S, Barizzone N, Ruzickova S, Dostal C, Schmidt RE, Witte T, Papasteriades C, Kappou-Rigatou I, Endreffy E, Kovacs A, Ordi-Ros J, Balada E, Carreira P, Gomez-Reino JJ, Gonzalez A: **Opposed independent effects and epistasis in the complex association of IRF5 to SLE.** *Genes Immun* 2007, **8**:429-438.
 18. Hochberg MC: **Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus.** *Arthritis Rheum* 1997, **40**:1725.
 19. Barrett JC, Fry B, Maller J, Daly MJ: **Haploview: analysis and visualization of LD and haplotype maps.** *Bioinformatics* 2005, **21**:263-265.
 20. Dupont WD, Plummer WJ Jr: **Power and sample size calculations. A review and computer program.** *Control Clin Trials* 1990, **11**:116-128.
 21. Todd JA: **Statistical false positive or true disease pathway?** *Nat Genet* 2006, **38**:731-733.
 22. Minegishi Y, Saito M, Morio T, Watanabe K, Agematsu K, Tsuchiya S, Takada H, Hara T, Kawamura N, Ariga T, Kaneko H, Kondo N, Tsuge I, Yachie A, Sakiyama Y, Iwata T, Bessho F, Ohishi T, Joh K, Imai K, Kogawa K, Shinohara M, Fujieda M, Wakiguchi H, Pasic S, Abinun M, Ochs HD, Renner ED, Jansson A, Belohradsky BH, et al.: **Human tyrosine kinase 2 deficiency reveals its requisite roles in multiple cytokine signals involved in innate and acquired immunity.** *Immunity* 2006, **25**:745-755.
 23. D'Esposito M, Quaderi NA, Ciccodicola A, Bruni P, Esposito T, D'Urso M, Brown SD: **Isolation, physical mapping, and northern analysis of the X-linked human gene encoding methyl CpG-binding protein, MECP2.** *Mamm Genome* 1996, **7**:533-535.
 24. Jacob CO, Zhu J, Armstrong DL, Yan M, Han J, Zhou XJ, Thomas JA, Reiff A, Myones BL, Ojwang JO, Kaufman KM, Klein-Gitelman M, McCurdy D, Wagner-Weiner L, Silverman E, Ziegler J, Kelly JA, Merrill JT, Harley JB, Ramsey-Goldman R, Vila LM, Bae SC, Vyse TJ, Gilkeson GS, Gaffney PM, Moser KL, Langefeld CD, Zidovetzki R, Mohan C: **Identification of IRAK1 as a risk gene with critical role in the pathogenesis of systemic lupus erythematosus.** *Proc Natl Acad Sci USA* 2009, **106**:6256-6261.
 25. Musone SL, Taylor KE, Lu TT, Nititham J, Ferreira RC, Ortmann W, Shifrin N, Petri MA, Kamboh MI, Manzi S, Seldin MF, Gregersen PK, Behrens TW, Ma A, Kwok PY, Criswell LA: **Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus.** *Nat Genet* 2008, **40**:1062-1064.
 26. Webb R, Wren JD, Jeffries M, Kelly JA, Kaufman KM, Tang Y, Frank MB, Merrill J, Kimberly RP, Edberg JC, Ramsey-Goldman R, Petri M, Reveille JD, Alarcón GS, Vilá LM, Alarcón-Riquelme ME, James JA, Vyse TJ, Moser KL, Gaffney PM, Gilkeson GS, Harley JB, Sawalha AH: **Variants within MECP2, a key transcription regulator, are associated with increased susceptibility to lupus and differential gene expression in patients with systemic lupus erythematosus.** *Arthritis Rheum* 2009, **60**:1076-84.

Supplementary Table 1
Origin and female percentage of the DNA sample collections used for replication of SLE associated SNPs.

Collections	DNA samples		Female %	
	Control	SLE	Control	SLE
Czech Republic ^a	99	101	32.3	85.1
Germany ^b	92	82	25	90.2
Greece 1 ^c	99	97	66.7	86.6
Greece 2 ^d	88	94	93.2	91.5
Hungary ^e	95	95	48.8	88.4
Italy, Milan ^f	106	128	42.5	86.7
Italy, Naples ^g	108	81	100	79
Italy, Rome ^h	102	84	55.9	83.3
The Netherlands ⁱ	180	104	58.9	86.5
Portugal ^j	95	94	89.5	85.1
Slovakia ^k	93	94	93.5	91.5
Spain, Asturias ^l	200	147	69	91.8
Spain, Barcelona ^m	96	89	52.1	91
Spain, Coruña ⁿ	81	88	71.6	92
Spain, Madrid ^o	98	92	64.3	91.3
Spain, Santiago ^p	95	109	48.4	89
Total	1726	1579	63.2	88.2

^a Institute of Rheumatology, Prague

^b Hannover Medical School

^c Evangelismos Hospital, Athens

^d Athens University Medical School

^e Albert Szent-Györgyi Medical and Pharmaceutical Centre, Szeged

^f University of Milan and Fondazione IRCCS Ospedale Maggiore

Policlinico

^g Second University of Naples

^h Ospedale S. Camillo – Forlanini, Rome

ⁱ University Medical Center Groningen

^j Hospital Garcia de Orta, Almada

^k Martin Faculty Hospital, Martin

^l Hospital Universitario Central de Asturias

^m Hospital Val d'Hebron, Barcelona

ⁿ CH Universitario Juan Canalejo, Coruña

^o Hospital 12 de Octubre, Madrid

^p Hospital Clínico Universitario de Santiago

Supplementary Table 2

Clinical characteristics of the patients with SLE included in the study

Clinical characteristic (scale)	Value
Age of disease onset (median, IQR)	28 (21-39)
Malar rash (%)	53.5
Discoid rash (%)	18.1
Photo-sensitivity (%)	53.5
Oral ulcers (%)	28.2
Arthritis (%)	80.7
Serosistis (%)	35.8
Renal disorder (%)	41.0
Neurologic disorder (%)	14.6
Hematologic disorder (%)	74.3
Immunologic disorder (%)	77.7
ANA (%)	93.0

Supplementary Table 3
Primers and probes in 5' to 3' direction used for genotyping the 10 SNPs by the SNaPshot technique

Locus	SNP	PCR primers		Minisequencing Probe
		Forward Primer	Reverse Primer	
<i>C8orf13-BLK</i>	rs13277113	cattaggtaaacctgtgcttttccct	ccaggtccaagattcacctcag	TTACCTATGATTGATCGTGGTGATATCCGctctgacgcaggcc atTTtaagatttaaacacttatcagatcatt \$
<i>ITGAM</i>	rs1143679	ccacaggggtggtgttgagc	cccaccacagatccccaggac	tacagcacagtctcatgcgagcccatcc
<i>KIAA1542</i>	rs4963128	ccgagatcttgactccgtccg	ggaggacaggaaggcaccaaag	TGATTGATCGTGGTGATATCCGcctgcctccccgagtgctgc c \$
<i>PXK</i>	rs6445975	gctaggatgtacctgccatttcc	gctgtccctagagccccagtc	actcacacttactgtaaagttaacaaaatggagca
<i>1q25.1</i>	rs10798269	cactaaggatggaccattttgagg	aggccatggctgggggaga	AGAACTATATTAAATGAGCAGTCAGcttgaggtcgaccaatct ctccctctttctcctctacaaatcta \$
<i>BANK</i>	rs17266594	gcaaatataatatatgacctagtgagc	acaggatggcttcccttttc	TGCTGGTGATATCCGctgcctagtgcacatattaaaatttttg agaaataataatttaacctgc \$
<i>MECP2</i>	rs17435	gccagtgagccagcccag	cggagagaatttgacgctg	ATCCGctgtagcttcagccctatctcggattctctctcttctgtt ctaagaactggttgg \$
<i>LY9</i>	rs509749	tcaactcaggaggcgctgg	tgccctgatccccctagtc	ACCTTCGGCGCAGATATAATtccatatgtcacggaagtgcag tctgtggttgagagaaacacc \$
<i>STA74</i>	rs7574865	tgggtgtgagtgagagtaagg	aatccccctgaaattccactga	catgagtggtgatgcagtaaaagtatgaaaagtgtggtgagca aaatgt
<i>TYK2</i>	rs2304256	ccctagcccagcccctac	aggaggtataaaacgggcattg	ccagcctgtttgggaaagaagcccaaggctcacaaggca

Primers selected with Primer3 and Oligos softwares. They were checked to avoid formation of dimers between the primers included in the reaction.
\$ These oligonucleotides were extended with a 5' tail that has no homology with human sequences (these tails are in capital letters)

Supplementary Table 4

Genotype counts for each of the ten SNPs and each of the sample collections
Data from sample collections that have been included in other publications were excluded (empty cells)
Data for the rs17205394 SNP from the Barcelona collection were excluded due to inconsistencies in genotyping

Sample collections	rs204256 (TYX2)			rs7574662 (STAT4)			rs644975 (PXK)			rs17432 (MECP2)			rs50749 (LTV9)			rs483126 (KOA1542)			rs143379 (ITGAM)			rs1327713 (C6orf13-3L4)			rs1726554 (BANK1)			rs10792623 (L225-1)		
	CC	CA	AA	GG	GT	TT	TT	TG	GG	AA	AT	TT	AA	AG	GG	GG	GA	AA	GG	GA	AA	GG	GA	AA	TT	TC	CC	GG	GA	AA
Portugal	57	32	5	44	35	15	40	42	12	96	30	2	27	44	23	46	30	18	52	36	3	50	40	4	57	36	1	47	41	6
Controls	46	36	13	52	40	3	51	36	8	61	32	2	29	42	24	51	35	9	61	31	3	56	36	3	42	45	8	38	46	11
SLE	66	23	3	39	46	7	55	30	7	96	34	2	34	44	14	57	26	9	57	29	6	47	34	11	53	36	3	63	25	4
Madrid	50	44	4	63	29	6	58	37	3	70	19	8	25	50	23	46	43	9	72	25	1	59	32	7	53	42	3	40	50	8
Controls	66	39	4	57	42	10	62	37	10	64	40	5	35	51	23	56	39	14	49	52	7	60	42	7	58	49	2	58	41	10
SLE	52	36	7	59	32	4	49	40	6	76	11	8	30	50	15	37	47	11	72	20	3	58	31	6	53	33	9	49	38	8
Santiago	77	64	6				80	60	7	96	44	7	40	70	37	67	64	16	82	56	9	81	50	16				78	62	7
Controls	112	75	13	134	61	5	138	43	18	50	104	45	50	104	45	85	83	32	138	57	5	123	67	10				86	86	28
SLE	47	35	6	47	33	8	48	38	2	43	34	10	24	46	18	52	30	6	54	27	7	42	40	6	55	29	4	52	34	2
Coruña	42	32	7	51	27	4	46	31	5	54	22	5	28	34	19	46	31	5	60	18	1	53	23	3	50	29	3	40	38	3
Controls	47	36	6	42	37	10	53	29	6	49	32	8	26	48	15	44	35	10	44	35	9	43	39	7				47	38	4
SLE	58	27	11	52	38	6	57	30	9	63	25	8	29	48	19	53	33	10	66	25	5	55	38	3				47	41	8
Barcelona	46	35	3				44	26	14				28	45	11	32	38	14	49	29	6							45	33	6
Controls	59	35	8	59	54	15	67	55	6	71	44	11	38	70	20	61	59	7	76	43	8	69	45	13	72	50	6	64	57	7
SLE	66	54	7	64	36	6	71	30	5				38	51	17	43	41	22	70	32	4	60	40	6	58	39	9	53	41	12
Milan	47	27	7							40	30	11	31	41	9	38	30	13	40	35	6							47	28	6
Controls	45	55	7							55	45	8	36	48	24	43	52	13	77	30	1							49	49	10
SLE	61	36	5				47	41	16	64	26	13	37	45	20	55	39	10	63	38	3	30	58	15	58	40	6	49	45	10
The Netherlands	94	71	15				92	73	15	127	29	21	65	88	27	90	78	12	130	48	2	63	95	22	98	68	14	74	62	24
Controls	57	35	3	48	39	8	46	39	10	47	38	10	32	51	12	52	33	10	64	30	1	45	38	12	46	38	11	48	44	3
SLE	69	25	3	58	35	2	46	46	3	67	17	10	28	38	29	37	45	13	77	17	1	55	35	5	48	39	8	42	42	11
Hungary	50	41	4	38	48	11	58	36	3	49	31	17	28	54	15	44	40	13	52	40	5	47	39	11	53	38	6	47	39	11
Controls	47	43	9	50	43	6	60	36	2	63	25	9	37	42	20	31	54	14	62	34	2	54	38	6	43	47	9	50	39	10
SLE	59	27	6	36	43	15	50	39	5	40	40	14	19	56	19	51	33	10	47	43	4	40	39	15	63	27	4	51	37	6
Greece 1	54	27	6	54	29	5	59	25	4	44	37	5	30	34	24	29	45	12	58	28	2	37	41	9	42	40	6	36	43	8
Controls	58	28	8	41	41	12	39	47	8	57	26	11	35	45	14	46	39	9	73	20	1	43	46	5	52	32	10	55	31	8
SLE	47	38	8	58	30	5	42	41	10	55	26	12	32	51	10	46	42	10	75	16	2	57	33	3	46	39	8	51	32	10
Greece 2	43	47	11	54	40	7	55	34	12	38	37	13	36	50	26	53	39	9	75	25	1	45	48	8	60	37	4	49	37	15
Controls	62	32	7	43	47	11	54	40	7	55	34	12	36	50	26	53	39	9	75	25	1	45	48	8	60	37	4	49	37	15
SLE	46	45	8	52	42	5	59	31	9	72	11	16	38	50	23	43	40	15	77	22	0	54	40	4	58	33	8	47	44	8
Czech Republic	45	31	6				44	33	4	47	30	3	29	46	7	41	31	7	53	24	5	38	32	12				32	41	9
Controls	45	39	8				51	33	8	66	22	4	32	40	20	49	31	11	75	17	0	49	33	10				45	39	8
SLE	930	559	87	494	465	122	743	566	103	878	539	155	501	793	283	795	605	175	930	562	84	680	590	142	627	412	57	832	633	114
Total	903	684	138	613	381	52	875	550	92	1150	402	165	559	821	346	767	741	216	1240	447	37	833	582	97	591	454	85	808	738	180

Supplementary Table 5

Minor allele frequencies for each of the ten SNPs and each of the sample collections
Data from sample collections that have been included in other publications were excluded (empty cells)
Data for the rs17266594 SNP from the Barcelona collection were excluded due to inconsistencies in genotyping

Sample collections		rs2304256 (TYK2)		rs7574865 (STAT4)		rs6445975 (PXX)		rs17435 (MECP2)		rs509749 (LV9)		rs4963128 (KIA1542)		rs1143679 (ITGAM)		rs13277113 (C6orf15-BLK)		rs17266594 (BANK1)		rs10798269 (10q5.1)	
		A	T	T	G	G	T	T	G	G	A	A	A	A	A	A	C	C	A	A	
Portugal	SLE	0.223	0.346	0.351	0.237	0.479	0.351	0.255	0.255	0.255	0.255	0.255	0.255	0.255	0.255	0.202	0.202	0.282	0.282		
	Controls	0.326	0.242	0.274	0.189	0.474	0.279	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.321	0.321	0.358	0.358		
Madrid	SLE	0.158	0.326	0.239	0.207	0.391	0.239	0.223	0.223	0.223	0.223	0.223	0.223	0.223	0.223	0.228	0.228	0.179	0.179		
	Controls	0.265	0.209	0.219	0.180	0.490	0.311	0.138	0.235	0.235	0.235	0.235	0.235	0.235	0.235	0.245	0.245	0.337	0.337		
Santiago	SLE	0.216	0.284	0.261	0.229	0.445	0.307	0.306	0.257	0.257	0.257	0.257	0.257	0.257	0.257	0.243	0.243	0.280	0.280		
	Controls	0.263	0.211	0.274	0.142	0.421	0.363	0.137	0.226	0.226	0.226	0.226	0.226	0.226	0.226	0.268	0.268	0.284	0.284		
Asturias	SLE	0.259	0.252	0.197	0.490	0.327	0.252	0.252	0.279	0.279	0.279	0.279	0.279	0.279	0.279	0.259	0.259	0.259	0.259		
	Controls	0.253	0.178	0.198	0.487	0.368	0.168	0.168	0.218	0.218	0.218	0.218	0.218	0.218	0.218	0.355	0.355	0.355	0.355		
Coruña	SLE	0.267	0.278	0.239	0.310	0.466	0.239	0.233	0.295	0.295	0.295	0.295	0.295	0.295	0.295	0.210	0.210	0.216	0.216		
	Controls	0.284	0.213	0.250	0.198	0.444	0.250	0.127	0.184	0.184	0.184	0.184	0.184	0.184	0.184	0.213	0.213	0.272	0.272		
Barcelona	SLE	0.270	0.320	0.233	0.270	0.438	0.309	0.301	0.298	0.298	0.298	0.298	0.298	0.298	0.298	0.258	0.258	0.258	0.258		
	Controls	0.255	0.260	0.250	0.214	0.448	0.276	0.182	0.229	0.229	0.229	0.229	0.229	0.229	0.229	0.297	0.297	0.297	0.297		
Rome	SLE	0.244	0.321	0.321	0.399	0.393	0.244	0.244	0.244	0.244	0.244	0.244	0.244	0.244	0.244	0.268	0.268	0.268	0.268		
	Controls	0.250	0.284	0.284	0.417	0.377	0.181	0.181	0.181	0.181	0.181	0.181	0.181	0.181	0.181	0.265	0.265	0.265	0.265		
Milan	SLE	0.268	0.328	0.262	0.262	0.430	0.287	0.232	0.232	0.232	0.232	0.232	0.232	0.232	0.232	0.242	0.242	0.277	0.277		
	Controls	0.283	0.226	0.189	0.198	0.401	0.401	0.189	0.245	0.245	0.245	0.245	0.245	0.245	0.245	0.269	0.269	0.307	0.307		
Naples	SLE	0.253	0.321	0.364	0.346	0.290	0.290	0.290	0.290	0.290	0.290	0.290	0.290	0.290	0.290	0.247	0.247	0.247	0.247		
	Controls	0.322	0.282	0.282	0.444	0.361	0.148	0.148	0.148	0.148	0.148	0.148	0.148	0.148	0.148	0.319	0.319	0.319	0.319		
The Netherlands	SLE	0.225	0.351	0.252	0.417	0.284	0.212	0.212	0.427	0.427	0.427	0.427	0.427	0.427	0.427	0.250	0.250	0.313	0.313		
	Controls	0.281	0.286	0.201	0.394	0.283	0.144	0.144	0.386	0.386	0.386	0.386	0.386	0.386	0.386	0.267	0.267	0.361	0.361		
Hungary	SLE	0.216	0.289	0.311	0.305	0.395	0.279	0.168	0.326	0.326	0.326	0.326	0.326	0.326	0.326	0.316	0.316	0.263	0.263		
	Controls	0.258	0.205	0.274	0.197	0.505	0.374	0.100	0.237	0.237	0.237	0.237	0.237	0.237	0.237	0.289	0.289	0.337	0.337		
Greece 1	SLE	0.160	0.361	0.216	0.335	0.433	0.340	0.258	0.314	0.314	0.314	0.314	0.314	0.314	0.314	0.258	0.258	0.314	0.314		
	Controls	0.308	0.278	0.204	0.222	0.414	0.194	0.194	0.255	0.255	0.255	0.255	0.255	0.255	0.255	0.328	0.328	0.298	0.298		
Greece 2	SLE	0.229	0.388	0.261	0.362	0.500	0.282	0.271	0.367	0.367	0.367	0.367	0.367	0.367	0.367	0.186	0.186	0.261	0.261		
	Controls	0.224	0.222	0.188	0.273	0.466	0.401	0.182	0.339	0.339	0.339	0.339	0.339	0.339	0.339	0.295	0.295	0.339	0.339		
Slovakia	SLE	0.234	0.346	0.335	0.255	0.388	0.303	0.117	0.298	0.298	0.298	0.298	0.298	0.298	0.298	0.277	0.277	0.250	0.250		
	Controls	0.290	0.215	0.328	0.269	0.382	0.333	0.108	0.210	0.210	0.210	0.210	0.210	0.210	0.210	0.296	0.296	0.280	0.280		
Czech Republic	SLE	0.228	0.342	0.267	0.287	0.441	0.282	0.134	0.317	0.317	0.317	0.317	0.317	0.317	0.317	0.223	0.223	0.332	0.332		
	Controls	0.308	0.263	0.247	0.217	0.384	0.357	0.111	0.245	0.245	0.245	0.245	0.245	0.245	0.245	0.247	0.247	0.303	0.303		
Germany	SLE	0.262	0.253	0.225	0.366	0.285	0.207	0.341	0.341	0.341	0.341	0.341	0.341	0.341	0.341	0.360	0.360	0.360	0.360		
	Controls	0.299	0.266	0.163	0.435	0.291	0.092	0.092	0.288	0.288	0.288	0.288	0.288	0.288	0.288	0.299	0.299	0.299	0.299		

Lack of Interaction Between Systemic Lupus Erythematosus-associated Polymorphisms in *TYK2* and *IRF5*

To the Editor:

These are exciting times for the genetic investigation of systemic lupus erythematosus (SLE), characterized by the discovery of many reproducibly associated loci¹. Further progress will require research in many different directions, including investigation of how the effects of each locus integrate between them and with environmental exposures to cause SLE. We read with interest the report by Hellquist, *et al*² that showed evidence of significant epistatic interaction between 2 SLE-associated loci, *IRF5* and *TYK2*. The first is a definitively confirmed SLE susceptibility locus with one of the strongest, albeit complex, effects. The latter has been more contentious, but its association with SLE is becoming clearer^{2,3}. Epistasis means that risk in subjects with susceptibility alleles at the 2 loci significantly exceeds the sum of the risks at each locus. This was rightly interpreted to mean that the 2 loci impinge in the type 1 interferon pathway², which is an important insight because *TYK2* codes for a Janus kinase that is involved in multiple cytokine signaling pathways in addition to this one⁴. Also, demonstration of epistasis between 2 SLE loci is of importance because its absence has been the rule for SLE genetic factors⁵ and for most genetic factors of other complex diseases^{6,7}. What is common is the contribution to total risk of the different disease alleles in a purely additive way, without any interaction. Therefore, we were interested in replicating the epistatic interaction described by Hellquist, *et al*² with the SLE data available to us.

We had already genotyped the relevant single-nucleotide polymorphisms (SNP), rs2304256 in *TYK2*³ and rs10954213 in *IRF5*⁸, in a large collection of patients with SLE and controls. These 2 SNP were associated with SLE in our studies ($p = 2.5 \times 10^{-5}$ and 0.015, respectively) with risk alleles that were coincident to those found by Hellquist, *et al* (who reported $p = 10^{-4}$ and $p = 0.004$, respectively). We have data for the 2 SNP in 419 SLE subjects and 454 controls. The same type of analysis done by Hellquist, *et al*², that is, a comparison between the fit to the data of logistic regression models with and without an interaction term, showed no differences (Table 1). As a more exhaustive and sensitive test, we also conducted analysis with the LRASSOC software⁹. This approach compares a range of specific genetic models with and without interaction and does not require statistical significance to discriminate between models, but only differences according to the less stringent Akaike's Information Criterion (AIC). The best model has the lowest AIC value, which means that it has the best fit to the case-control genotypes with the highest combination of likelihood and parsimony. These analyses indicated that the best model included the independent contribution of the 2 loci without any interaction, although models with interaction terms were not much worse (Table 1). Therefore, we did not confirm the epistatic interaction that Hellquist and colleagues have described.

As a complementary analysis, we checked whether there was statistical evidence of interaction between *TYK2* and *IRF5* using other SNP in *IRF5*. These SNP were rs10488631 (which in some studies¹⁰ has been replaced

by its proxy, rs2070197) and rs729302. The first SNP showed the strongest association ($p = 4.8 \times 10^{-20}$) with SLE in our study⁸ and others¹⁰ and accounts for association of many other SNP in the gene, including rs10954213^{8,10}. The second, rs729302, is an SNP in the promoter region of *IRF5* that was independently associated to SLE in our samples ($p = 1.6 \times 10^{-9}$). Neither of the 2 SNP showed evidence of interaction with the rs2304256 SNP of *TYK2*, either with an analysis as done by Hellquist, *et al*² or with LRASSOC (Table 1). These comparisons are more powerful than models with rs10954213 because the 2 *IRF5* SNP are much more strongly associated to SLE susceptibility and because we have data for 3-fold more samples: 1223 cases and 1300 controls.

Our results did not show any statistical evidence of interaction between SLE-associated SNP in *TYK2* and *IRF5*. As a consequence, we should continue to consider that these 2 SLE genetic factors can contribute to disease susceptibility by independent pathways, unless other genetic studies replicate the epistatic interaction or functional studies support the interaction. Similar interaction analysis for other genetic factors and in other sample collections will unravel the influence of epistatic interaction in SLE susceptibility, but current evidence does not support an important contribution⁵ (Suarez-Gestal, *et al*, unpublished data).

MARIAN SUAREZ-GESTAL, MSc; MANUEL CALAZA, MMath; ANTONIO GONZALEZ, MD, PhD; European Consortium of SLE DNA Collections, Laboratorio de Investigacion 10 and Rheumatology Unit, Hospital Clínico Universitario de Santiago, Santiago de Compostela, Spain. Address correspondence to Dr. Gonzalez; E-mail: Antonio.Gonzalez.Martinez-Pedrayo@sergas.es

REFERENCES

- Rhodes B, Vyse TJ. The genetics of SLE: an update in the light of genome-wide association studies. *Rheumatology* 2008;47:1603-11.
- Hellquist A, Järvinen TM, Koskenmies S, Zucchelli M, Orsmark-Pietras C, Berglind L, et al. Evidence for genetic association and interaction between the *TYK2* and *IRF5* genes in systemic lupus erythematosus. *J Rheumatol* 2009;36:1631-8.
- Suarez-Gestal M, Calaza M, Endreffy E, Pullmann R, Ordi-Ros J, Domenico Sebastiani G, et al. Replication of recently identified systemic lupus erythematosus genetic associations: a case-control study. *Arthritis Res Ther* 2009;11:R69.
- Minegishi Y, Saito M, Morio T, Watanabe K, Agematsu K, Tsuchiya S, et al. Human tyrosine kinase 2 deficiency reveals its requisite roles in multiple cytokine signals involved in innate and acquired immunity. *Immunity* 2006;25:745-55.
- International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEn), Harley JB, Alarcón-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, Moser KL, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXX, KIAA1542 and other loci. *Nat Genet* 2008;40:204-10.
- Wellcome Trust Case Control Consortium. Genome-wide

Table 1. Lack of epistatic interaction between the SLE-associated SNP in *TYK2* and *IRF5*.

	SNP		No. of Subjects		Models with or without Interaction*		LRASSOC Models **	
	<i>TYK2</i>	<i>IRF5</i>	SLE	Controls	Best Model	p	Best Model	AIC to Interaction
Hellquist ²	rs2304256	rs10954213	277	356	With interaction	0.014	—	—
Our data ^{3,8}	rs2304256	rs10954213	419	454	No difference	0.68	Additive†	1.83
	rs2304256	rs10488631	1223	1300	No difference	0.3	Additive	1.98
	rs2304256	rs729302	1223	1300	No difference	0.88	Additive	0.94

* Multivariate logistic regression models with and without an interaction term were compared as in Hellquist, *et al*². The models did not specify any mode of inheritance. P values were obtained with the likelihood ratio test. ** Comparison of the specific inheritance models included in LRASSOC⁹. Akaike's Information Criterion (AIC) was used to classify the models. A value of AIC > 2 indicates a meaningful difference. † Additive: model with additive effects of each SNP, without dominance or interactive components. SNP: single-nucleotide polymorphisms.

- association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661-78.
7. Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, Goldgar D, et al. Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappa-B pathways. *Nat Genet* 2009;41:199-204.
 8. Ferreiro-Neira I, Calaza M, Alonso-Perez E, Marchini M, Scorza R, Sebastiani GD, et al. Opposed independent effects and epistasis in the complex association of IRF5 to SLE. *Genes Immun* 2007;8:429-38.
 9. North BV, Curtis D, Sham PC. Application of logistic regression to case-control association studies involving two causative loci. *Hum Hered* 2005;59:79-87.
 10. Kozyrev SV, Lewén S, Reddy PM, Pons-Estel B, Argentine Collaborative Group, Witte T, et al. Structural insertion/deletion variation in IRF5 is associated with a risk haplotype and defines the precise IRF5 isoforms expressed in systemic lupus erythematosus. *Arthritis Rheum* 2007;56:1234-41.
- J Rheumatol* 2010;37:3; doi:10.3899/jrheum.090823

2. Estudio funcional de PD1.3



ORIGINAL ARTICLE

Analysis of the functional relevance of a putative regulatory SNP of PDCD1, PD1.3, associated with systemic lupus erythematosus

M Suarez-Gestal^{1,4}, I Ferreiros-Vidal^{1,4}, JA Ortiz², JJ Gomez-Reino^{1,3} and A Gonzalez¹

¹Laboratorio de Investigación 2 and Rheumatology Unit, Hospital Clínico Universitario de Santiago, Santiago de Compostela, Spain;

²Instituto Bernabeu, Departamento de Biología Molecular, Alicante, Spain and ³Department of Medicine, University of Santiago de Compostela, Santiago de Compostela, Spain

This study aimed to test the functional effects of the PD1.3 single nucleotide polymorphism (SNP) (rs11568821), which were proposed based on its association to systemic lupus erythematosus (SLE) susceptibility and in electrophoretic mobility shift assays (EMSA) results. We analysed transcriptional effects of the PD1.3 locus by enhancer reporter assays. Results were against the hypothesis that the PD1.3 locus acts as enhancer in transcriptional regulation of PDCD1. In addition, they excluded a differential effect of the PD1.3 alleles. EMSA results confirmed that oligonucleotides with the PD1.3 G allele bind RUNX1 but not those with the A allele. However, binding to PD1.3 G oligonucleotides was much lower than binding to positive control oligonucleotides. Criss-cross experiments showed that this was due to flanking nucleotides in the PD1.3 sequence that negatively affect RUNX1 binding. These results cast doubts on the functional relevance of the PD1.3 SNP and, together with the lack of association in several studies, put into question its role as an SLE susceptibility factor. Investigation of other PDCD1 polymorphisms is needed to uncover the possible effect of this gene on SLE susceptibility.

Genes and Immunity advance online publication, 10 April 2008; doi:10.1038/gene.2008.19

Keywords: systemic lupus erythematosus; functional polymorphism; transcriptional regulation; transcription factor binding site

Introduction

Investigation of the genetic component of complex diseases involves both genetic epidemiology studies and functional tests. Reports with convincing epidemiological and functional evidence are regarded as especially robust. One such report by Prokunina *et al.*¹ has identified a regulatory polymorphism in PDCD1 intron 4, called PD1.3, as a susceptibility factor for systemic lupus erythematosus (SLE; [MIM] 152700). The finding was supported by association in a large number of samples from different ethnic groups and by differential allele binding of AML1/RUNX1 transcription factor in electrophoretic mobility shift assays (EMSA).

Association of PD1.3 with SLE has been replicated in some studies.^{2–3} However, most posterior studies have failed to replicate the PD1.3-SLE genetic association. In some studies, there was not association,^{4,5} in others association was only with specific subsets of SLE patients, or association was with other polymorphisms in the

PDCD1 gene^{3,6–8} or with the opposite allele of the PD1.3 single nucleotide polymorphism (SNP).⁹ These contradictory results could be related to an allele frequency cline in PD1.3, which introduces a bias in association studies depending on the population explored.⁵

As a consequence of the discordant results, it is still unclear if genetic variation in PDCD1, either PD1.3 or another polymorphism, plays any role in SLE susceptibility. This is very relevant because SLE has a clear genetic component, recurrence sibling ratio of 20–40, and few genetic factors identified.^{10,11} In addition, PDCD1 is a critical regulator of acquired immunity, and relevant genetic variants could have repercussion in other diseases.¹² It has two characteristic tyrosine domains in the cytoplasmic tail: an immunoreceptor tyrosine-based inhibition motif and an immunoreceptor tyrosine-based switch motif (ITSM). Interaction of PDCD1 with its widely expressed ligands (PD-L1 and PD-L2) leads to tyrosine-phosphorylation of immunoreceptor tyrosine-based inhibition motif (ITIM) and immunoreceptor tyrosine-based switch motif domains producing down-regulation of B- and T-cell responses. Experimental animal models have established that PDCD1 is required for the avoidance of autoimmunity, deficient animals showing exacerbated autoimmunity. In addition, PDCD1 signals are the cause of persistent viral infections by downregulating T CD8 reactivity, these signals can also be involved in inefficient immune responses against

Correspondence: Dr A Gonzalez, Laboratorio de Investigación 2, Hospital Clínico Universitario de Santiago, Travesía de Choupana sn., Santiago de Compostela 15706, Spain.

E-mails: Antonio.Gonzalez.Martinez.Pedrayo@sergas.es or anlugon@hotmail.com

⁴These authors contributed equally to this work.

Received 28 December 2007; revised 28 February 2008; accepted 28 February 2008

tumours and have been found to modulate atherogenesis in the LDLR knockout model.¹³ The involvement of PDCD1 in chronic viral infections has already been confirmed in human diseases of great impact such as AIDS¹⁴ and chronic viral hepatitis.¹⁵ Therefore, if genetic variation of PDCD1 is associated with changes in PDCD1 function, it could have repercussions in human health well beyond SLE.

A point that has not been explored after the Prokunina *et al.*¹ report is the proposed regulatory role of the PD1.3 SNP. This SNP is in a region of PDCD1 intron 4 that was described as enhancer-like due to the presence of four tandem repeats that contain multiple putative binding sequences of transcription factors. The A allele of PD1.3 destroys a putative binding site of RUNX1. RUNX1 consensus binding site, TG(T/C)GGT, was described by Meyers *et al.*,¹⁶ who analysed random oligonucleotides selected for affinity to RUNX1. This binding site is functional in multiple cellular enhancers and promoters of genes implicated in hematopoiesis and immune system function, such as TCR, IL-3, GM-CSF, CD3 ϵ , myeloperoxidase 1 α , IL-2 and INF γ .¹⁷

Herein, we aimed to fully explore the function of the PD1.3 SNP. As a first test we evaluated in enhancer reporter assays the repetitive structure in PDCD1 intron 4 and whether there was any differential effect of the two PD1.3 alleles in regulating gene expression. The negative results in these experiments led us to replicate EMSA. We observed that RUNX1 binding to the PD1.3 G allele sequence was stronger than to the A allele, but much weaker than to positive control binding sites. These results put into question the functional relevance of this SNP.

Materials and methods

Analysis of tandem repeats and orthologous sequences

The tandem repeats finder program integrated in the USCS genome browser¹⁸ was used to localize repeats in PDCD1 intron.¹⁹ The Ensembl release 46 from August 2007²⁰ and the USCS genome browser were used for the analysis of PDCD1 orthologous sequences. The Phast-Cons program of the USCS browser was used to obtain an estimate of overall conservation across vertebrates and placental mammals.²¹ In Ensembl, only placental mammalian sequences were analysed. Sequences of the region surrounding PD1.3 SNP from all species with orthologous genes were extracted and aligned.

Enhancer luciferase reporter vectors

Genomic DNA from a PD1.3 heterozygous subject was used to generate 514bp PCR fragments containing the enhancer-like structure of intron 4 of PDCD1 gene with the two alleles. PCR was done in 10 μ l containing 30 ng of DNA, 0.1 μ M of each primer (forward primer 5'-TGGAAGGACAGGCTGGGACC-3' and reverse primer 5'-AGTACCGGACCGAACCCTGG-3'), 1.5 mM of MgSO₄, 200 nM of dNTPs, 1 U of Easy-A PCR cloning enzyme (Stratagene, La Jolla, CA, USA) and 1 \times reaction buffer (PCR Enhancer System, Invitrogen, Paisley, UK). PCR conditions were: denaturation at 95 °C for 3 min followed by 35 cycles of denaturation at 95 °C for 15 s, annealing at 56 °C for 15 s and extension at 72 °C for 15 s. Amplified fragments were cloned into the EcoRV site of pBluescript (pBK SK⁻) vector by TA cloning. The EcoRV

(New England Biolabs, Ipswich, MA, USA) cut vector was treated with Taq DNA polymerase (Roche, Sant Cugat del Vallés, Spain) in the presence of 2 mM of dTTP and PCR buffer to add thymine to its 3' ends. TA cloning was done in a 10 μ l reaction containing 50 μ g of the 3' T-vector, 25 μ g of PD1.3 PCR product, 1 μ l of T4 DNA ligase (New England Biolabs) and 1 \times ligase buffer at 16 °C overnight. Cloned PD1.3 fragments were subcloned into the *KpnI/SmaI* (upstream) or *BamHI/SalI* (downstream) sites of a *fos*-pGL3-basic vector. This vector contains the minimal *fos* promoter upstream of the *Firefly* luciferase reporter gene. We generated eight reporter vectors, four with each PD1.3 allele, two upstream and two downstream of the luciferase gene, and one of each pair in a different orientation: sense and antisense (Figures 2 and 3). Sequence and orientation of each insert was verified by sequencing.

Reporter assays

We transfected 10⁶ Jurkat cells in 0.8 ml of OptiMEN medium (Invitrogen) with 0.8 μ g of *fos*-pGL3luc vectors, either without insert or with any of the eight 'enhancer' inserts, and 0.2 μ g of pRL-TK *Renilla* luciferase vector (Promega, Madison, WI, USA) by using 2 μ l of Lipofectamine (Invitrogen) and 3 μ l of Plus reagent (Invitrogen). Some wells were stimulated the following day with 500 ng ml⁻¹ of ionomycin and 10 ng ml⁻¹ of phorbol 12-myristate 13-acetate (PMA). After 4 h, cells were harvested and luciferase activity was measured using the *Firefly* Luciferase Assay System and the *Renilla* Luciferase Assay System (both of Promega). Results of *Firefly* luciferase were expressed as relative luciferase units after correction by *Renilla* luciferase activity to adjust for transfection efficiency. Comparison of expression levels was done with the Wilcoxon matched pairs test.

Nuclear extracts

For preparation of nuclear extracts, 10⁷ Jurkat or MOLT-4 cells were washed in 1 ml of 1 \times phosphate-buffered saline and lysed at 4 °C for 15 min in 200 μ l of a buffer containing 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid pH 7.9, 1 mM ethylenediaminetetra acetic acid (EDTA), 1 mM ethylene glycol bis(β -aminoethyl)-N,N,N',N'-tetraacetic acid, 10 mM KCl, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, 10 μ g ml⁻¹ aprotinin and 10 μ g ml⁻¹ leupeptin (all from Sigma-Aldrich, Madrid, Spain). Nuclei were incubated at 4 °C for 10 min with 0.1% of Triton-X100 and collected by centrifugation at 800 \times g and 4 °C for 15 min. Nuclei were lysed at 4 °C for 30 min with rotation in 200 μ l of the same buffer used to lyse cells plus 20% glycerol and 0.4 M KCl. Insoluble material was precipitated by centrifugation at 13000 r.p.m. at 4 °C for 15 min and proteins were quantified by Bradford's method.

Electrophoretic mobility shift assays

Double-stranded biotin-labelled oligonucleotides with the PDCD1 sequence of PD1.3 and with two other sequences with consensus binding sites for AML1/RUNX1 (Table 1) were synthesized by Sigma-Genosys (Cambridge, UK). Each double-stranded oligonucleotide was annealed at 37 °C for 1 h, and 0.5 mM were incubated with Jurkat or MOLT-4 cell nuclear extracts in 20 μ l of binding buffer (10 mM Tris-HCl pH 7.5, 90 mM KCl,

1.4 mM dithiothreitol, 50 ng μL^{-1} poly(dI·dC), 0.1 mM EDTA y 1 mM MgCl_2 for 20 min at room temperature. We also tested other binding conditions^{1,16} to check if the weak signals we obtained with PDCD1 oligonucleotides were protocol-dependent. For competition experiments, unlabelled oligonucleotides were added to the nuclear extracts at a 200-fold molar excess (0.1 μM) before the addition of the biotin-labelled probe. For supershift experiments, nuclear extracts were incubated with 2 μg of polyclonal antiserum against the N-terminal region of AML-1 (sc-8563, Santa Cruz Biotechnology, Santa Cruz, CA, USA) at room temperature for 20 min before the addition of oligonucleotides. DNA-protein complexes were analysed by electrophoresis at 11 V cm^{-1} and 4 °C in 5% polyacrylamide gels and 0.5 \times TBE. Complexes were transferred to a nylon membrane at 12 V for 45 min, and crosslinked to the membrane at 7.2 J cm^{-2} using a transilluminator. Migration of the biotinylated oligonucleotides and their complexes was detected by chemiluminescence using the LightShift Chemiluminescence Detection Module (Pierce, Rockford, IL, USA) followed by exposure of the membrane to X-ray films.

Results

Lack of conservation of the putative enhancer in PDCD1 intron 4

The putative enhancer in the 4th PDCD1 intron has been proposed based on a four tandem repeat structure

Table 1 Oligonucleotides used in EMSA

Symbol	Gene	Sequences ^a
P ₁₂	PDCD1_PD1.3G	ACCTGCGGTCTC
P ₁₂ ^m		AttTGGCGGTaC
P _{18G}	PDCD1_PD1.3G	CCACCTGCGGTCTCCGG
P _{18A}	PDCD1_PD1.3A	CCCACCTGCACTCTCCGG
C ₁	Consensus	GGATATTGCGGTAGCA
C ₁ ^m		GGATAccTGGCGTcGCA
C ₂	Consensus	AATTCGAGTATTGGTTAATACG
		GCTCATAACCAATTATGCTTAA

Abbreviations: EMSA, electrophoretic mobility shift assays; SNP, single nucleotide polymorphism.

Double-stranded oligonucleotides were used with exactly complementary sequences except for C₂ where the two complementary oligonucleotides are shown. Sense and antisense oligonucleotides were biotin-labelled at their 5' end.

^aConsensus hexanucleotide binding site is underlined; SNP nucleotide is in bold; modified nucleotides are in lower case.

including multiple transcription factor binding sites (TFBS).¹ Analysis with the tandem repeats finder program¹⁹ integrated in the UCSC genome browser showed two alternative and overlapping tandem repeat structures (Figure 1): one with four repeats of 40 bp with an 83% identity and without insertion/deletions that included PD1.3 in its most telomeric repeat, and a second structure of two repeats of 80 bp with 90% identity and absence of insertion/deletions. These 80 bp repeats did not overlap completely with the four repeats of 40 bp and the most telomeric of them finished just before the PD1.3 SNP.

As functional elements are more conserved between species than non-functional sequence regions, we tested if the area containing the repeat sequences was evolutionary conserved. First we looked for the overall conservation of 400 bp surrounding PD1.3 and including the tandem repeats among vertebrates in the UCSC genome browser.²¹ The PhastCons program provided no evidence of conservation across 28 vertebrate species or across 17 placental mammals.

In addition, we searched for PDCD1 orthologous genes in the Ensembl genome database²⁰ and found that they were present in six other species of placental mammals. The PDCD1 sequence in *Macaca mulatta* showed a large sequence gap including the whole 4th intron and was excluded of further analysis. The other five genomic sequences included orthologous regions both upstream and downstream of the region of interest. However, alignment of the region surrounding PD1.3 was poor or impossible for *Canis familiaris* and *Bos torus*. Of the three sequences remaining, those from *Mus musculus* and *Rattus norvegicus* showed important lack of alignment in relation with the human sequence (Figure 1). Only the genomic sequence from *Pan troglodytes* showed a good alignment to the PD1.3 neighbouring sequences (T allele at the PD1.3 SNP). These alignments showed that humans and chimpanzees share exactly one of the two 80 bp repeats, the most telomeric, whereas the other is absent in chimpanzees. This suggests that the 80 bp repeat is the element that has been duplicated in the human genome from an original single copy present in a primate ancestor. Therefore, the putative enhancer in these repeats is a specific human feature with a single copy antecedent in chimpanzees.

Gene reporter assays exploring the enhancer effect

Enhancers have been defined by their effect increasing gene transcription in a manner that is largely position- and orientation-independent. To test the function of the

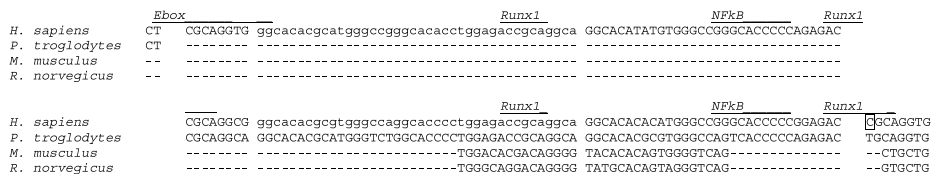


Figure 1 Alignment of the tandem repeat structure near the PD1.3 SNP (encased) in PDCD1 intron 4 with orthologue mammalian genome sequences. The plus strand is presented although the PDCD1 gene is transcribed in the minus strand (G and A alleles in place of the C and T presented here). Spaces are adjusted to highlight tandem repeat structure: two 80 bp repeats and partly overlapping four 40 bp repeats. The latter are indicated by alternating lower and upper case and separated by spaces. Some potential transcription factor binding sites in the minus strand of PDCD1 have been signalled.

putative enhancer in the repetitive sequence of PDCD1 intron 4, we transiently transfected Jurkat T-cell lines with luciferase reporter gene constructs containing this structure. The fragment of the 4th intron that was cloned contained the two 80bp repeat sequences, including the G allele of PD1.3 (the common allele consistent with the RUNX1 consensus binding site) and 386 nucleotides centromeric and 127 nucleotides telomeric to it (514 bp in total). This fragment was inserted in a vector containing the *Firefly* luciferase gene under the control of a *fos* minimal promoter. Cloning sites were selected to introduce the 4th intron fragment both 5' and 3' to the reporter gene and in sense and antisense orientations (Figure 2). In this way we expected to detect enhancer effects and to be able to differentiate them from potential artefacts. Five independent experiments were performed. They showed lack of increase in luciferase expression in cells transfected with any of the 4th intron containing constructs in relation with the empty *fos-luc* vector (Figure 2a for basal conditions and Figure 2b for post-stimulation). There was a trend to lower expression in the cells transfected with the constructs placed upstream of the *fos* promoter, especially after PMA + ionomycin stimulation where it became significant ($P = 0.043$). This restricted downmodulation suggested a positional effect (Figure 2b).

Lack of evidence of PD1.3 allele effect on reporter gene transcription

We used the same enhancer reporter gene system to compare the effect on gene expression of inserts containing either the major G allele or the minor A allele of the PD1.3 SNP. Results from five experiments showed that luciferase expression was equivalent with the two PD1.3 alleles, both in basal (Figure 3a) and in stimulated cells (Figure 3b) and with independence of the upstream or downstream position of the insert in relation with the reporter gene or of its orientation.

Inefficient binding of RUNX1 to the PD1.3 site

Given the lack of differential effect of the PD1.3 alleles in reporter gene assays, we decided to replicate the EMSA that supported the functional relevance of the PD1.3 SNP. An oligonucleotide of the same length as reported,¹ 18 nucleotides, and containing the PD1.3 G allele in a centre position, P_{18G} (Table 1), gave very weak signals in EMSA with Jurkat or MOLT-4 nuclear extracts. By increasing the amount of nuclear extracts and extending exposure times, we were able to replicate the specific binding with the PD1.3 G allele oligonucleotide, P_{18G} , and lack of binding with the A allele oligonucleotide, P_{18A} (Figure 4). Bound transcription factor was AML1/RUNX1 as shown by supershift with specific antibodies (Figure 5). However, the weak signal obtained with the G allele

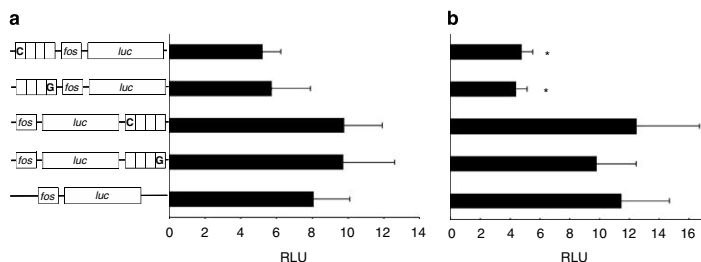


Figure 2 Luciferase enhancer reporter assays exploring the tandem repeat region of PDCD1 intron 4 containing the PD1.3 SNP. Vectors containing this region in the forward or reverse orientation and upstream or downstream of a *Firefly* Luciferase gene (*luc*) under control of a minimal *fos* promoter, were compared with the activity of a vector without it in Jurkat cells (a) in basal conditions or (b) after phorbol 12-myristate 13-acetate and ionomycin stimulation. These graphs represent the mean and s.e. of five independent experiments. * $P < 0.05$.

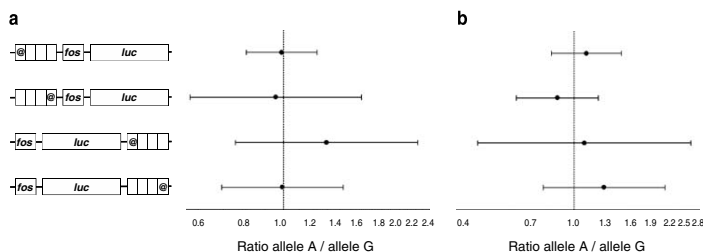


Figure 3 Lack of difference between the G and A allele at the PD1.3 SNP in luciferase enhancer reporter assays. The two PD1.3 alleles were compared in transiently transfected Jurkat cells: (a) in basal conditions and (b) after phorbol 12-myristate 13-acetate and ionomycin stimulation. Vectors were as in Figure 2 but either with the G or with the A allele at PD1.3 (their position represented by @). Results are geometric means and 95% confidence interval of the ratio between luciferase signal in cells transfected with the A allele and cells transfected with the G allele in homologous vectors. Five independent experiments were analysed.

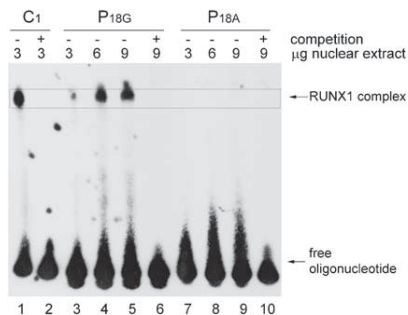


Figure 4 Specific but weak binding of Jurkat nuclear extracts to an 18bp oligonucleotide containing the PD1.3 G allele, P_{18G} (lanes 3–6). No binding was observed with the PD1.3 A allele oligonucleotide P_{18A} (lanes 7–10). Specificity of P_{18G} binding was shown by competition with unlabelled P_{18G} (lane 6). A control 18bp oligonucleotide, C₁, with the same RUNX1 consensus binding hexanucleotide showed much stronger binding at the same nuclear extract concentration (lane 1 versus 3). The need to use large amounts of nuclear extracts and long exposure times to visualize P_{18G} binding led to increased background and unspecific bands, and to saturation of the stronger bands that prevented quantitative comparisons.

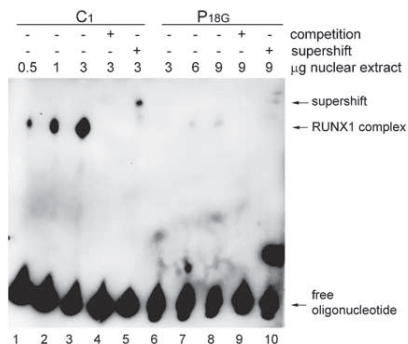


Figure 5 Comparison of RUNX1 binding to the intron 4 PDCD1 oligonucleotide P_{18G} with binding to the control oligonucleotide C₁. These two oligonucleotides share the same RUNX1 consensus hexanucleotide. Competition with unlabelled oligonucleotides (lanes 4 and 9) showed the specificity of binding and supershift with anti-AML1 antibody identified the bound factor (lanes 5 and 10). Differences in binding signals varied between 300- and 4000-fold.

oligonucleotide P_{18G} prompted us to compare this signal to the obtained with sequences known to be functional RUNX1 binding sites. We selected two (Table 1), C₁ with the same hexanucleotide sequence in the consensus binding site (TGCGGT) as PD1.3 allele G and also of 18 nucleotides.²² A second oligonucleotide, C₂, defined in the study of the RUNX1 consensus binding site, includes an hexanucleotide sequence that differs in the third base

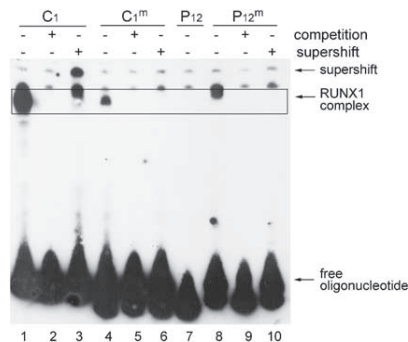


Figure 6 Effect of nucleotides flanking the RUNX1 consensus binding hexanucleotide. Electrophoretic mobility shift assays comparing binding of unmodified control oligonucleotide C₁ to binding to a C₁ modified oligonucleotide, C₁^m that contains the four nucleotides, two on each side, flanking the central hexanucleotide from the PD1.3 site. It also compares binding to a PD1.3 oligonucleotide, P₁₂, and its modified version, P₁₂^m containing the four flanking nucleotides from C₁. Binding was much lower to C₁^m (lane 4) than to C₁ (lane 1), and much stronger to P₁₂ (lane 8) than P₁₂^m (lane 7) showing the important role of flanking nucleotides. DNA-protein complexes were specific, as shown by competition with unlabelled oligonucleotides (lanes 2, 5 and 9), and the protein was identified as RUNX1 by supershift with a specific antibody (lanes 3, 6 and 10).

(TGTGGT).¹⁶ The two control oligonucleotides produced much stronger signals than the P_{18G} sequence (results of C₂ binding are not shown). This difference was observed with Jurkat and MOLT-4 nuclear extracts and both with extracts obtained in basal conditions or after stimulation with PMA + ionomycin (not shown). The low binding of the P_{18G} oligonucleotide was independent of the technique, as was observed with three different binding protocols (not shown). We quantified the difference in signal by serial dilution of nuclear extracts and densitometry reading (Figure 5). Direct comparison between P_{18G} and C₁ oligonucleotides showed, in different experiments, a 300- to 4000-fold lower intensity of DNA complexes with the P_{18G} oligonucleotide.

These results provoked us to review the reports defining the AML1/RUNX1 binding site. It was apparent that nucleotides flanking the central hexanucleotide could be important. We tested this hypothesis by designing modified oligonucleotides exchanging the four flanking nucleotides, two at each side of the hexanucleotide, between the RUNX1 binding site in C₁ and the putative site around PD1.3 (Table 1). According to our hypothesis, the modified control, C₁^m oligonucleotide, showed much weaker RUNX1 binding than the original C₁ oligonucleotide (Figure 6). On the contrary, the modified PD1.3 oligonucleotide, P₁₂^m showed a clear binding signal, whereas no binding was ever detected with the P₁₂ oligonucleotide (Figure 6). Binding to C₁, C₁^m and P₁₂^m were shown to be due to RUNX1 by supershift after antibody incubation (Figure 6). Therefore, the PDCD1 sequence surrounding PD1.3 bound RUNX1 very inefficiently due to the nucleotides flanking the described hexanucleotide consensus site.

Discussion

Identification of genetic variation in regulatory elements of the PDCD1 gene is of great interest given the critical role of this immunoregulatory receptor in multiple diseases.^{1,12,14,15} In some of them, PDCD1 downregulation of immune reactions is defective, as in autoimmunity and, possibly, in atherosclerosis,¹³ whereas in others it is inappropriately strong, as in chronic viral infections. The hypothesis of an enhancer in the 4th intron of the gene that bears a regulatory polymorphism deserves, therefore, to be tested.

Analysis of the sequence by bioinformatic means can only provide suggestive information. The enrichment in several TFBS related with immune function has been taken as supportive of the proposed enhancer.¹ However, bioinformatic identification of TFBS has low predictive value. This is due to the lack of knowledge of the factors in the DNA sequence that influence TF binding to DNA and to the effect of chromatin modifications independent from the DNA sequence (as DNA methylation, nucleosome position and histone modifications). In addition, TFBS are predicted on the basis of studies of very few validated sequences and the defined consensus oligonucleotides and nucleotide matrices often are very short, that is, hexanucleotides, and allow for substitutions at more than one place. These characteristics determine a low specificity in the identification of TFBS.²³ For example, we searched for binding sites for RUNX1 over the PDCD1 gene in TRANSFAC Professional/Match (BioBase Biological Databases) and found 70, only 11 in the 4th intron.

Transcription factor binding site prediction will improve in the next few years thanks to more complex bioinformatic tools and to a better knowledge of functional TFBS.²⁴ For example, whole genome studies with chromatin immunoprecipitation have allowed a better definition of consensus binding sites that have a much stronger predictive value.^{25,26} Often these consensus sequences are much larger than the previously described or include influence from neighbouring nucleotides. Unfortunately, there are only a few of these studies and they do not include RUNX1.

The presence of tandem repeats of several decenes of nucleotides has been observed as part of some complex regulatory units, including several enhancers.^{27–29} Some of them are polymorphic and affect transcriptional regulation, but in all the described examples the functional polymorphism affects the number of copies of the repeat. It is less likely that an SNP affecting a single TFBS in a repeated structure will have functional consequences, as it will be compensated by the same TF bound to the other repeats.

Another factor that is possible to analyse is sequence conservation. Functional sequences are conserved through evolution more often than neutral sequences.²¹ Neither the tandem repeats nor the PD1.3 sequence were found to be conserved even among mammals. However, this does not contradict functional relevance as only a fraction of the functional elements are conserved.

Experimental testing of PDCD1 intron 4th structures in a classical enhancer reporter assay showed no evidence of an enhancer effect on the expression of the luciferase gene. There was only a slight but significant decrease in expression after stimulation in the constructs with the

intron 4th inserted upstream to the reporter gene. This could be caused by unspecific interference with the minimal promoter but not to an enhancer effect because enhancers are largely independent of position and orientation. More important for the hypothesis being tested was the lack of difference in reporter gene expression when the two PD1.3 alleles were compared.

These negative results prompted us to re-evaluate the differential binding of the PD1.3 alleles to RUNX1 in EMSA. We replicated the preferential binding of the G allele but it was of much lower intensity than positive control oligonucleotides. This difference was explained by the lack of optimal flanking nucleotides in the PDCD1 sequence as demonstrated by criss-cross experiments.

Inefficient binding of RUNX1 to the PD1.3 sequence was overlooked by the TFBS searching softwares that only consider the central hexanucleotide TG[*T*/c]GGT. However, Thornell *et al.*³⁰ had already seen evidence that other nucleotides around the consensus sequence could interfere in RUNX1 binding to DNA. They suggested a decamer TTTGCGGTGA/T as the consensus binding site. This is the same sequence present in the two oligonucleotides with which we observed higher binding in our assays.

Our findings are of special relevance in the analysis of the role of genetic variation of PDCD1 in human disease. The PD1.3 SNP has had a dominant role in the genetic epidemiology studies of PDCD1 because it was found associated to SLE in a very impressive association study and because it was supposed to be functional.¹ However, most subsequent reports have not found the reported genetic association and an explanation of these discordant results, at least among Europeans, has been proposed⁵ (Asian populations are not polymorphic at this SNP³¹). Now, the proposed functional effects of PD1.3 are also questioned by our results. It is still possible that this SNP could be functional in ways that we have not explored, but available information does not allow predicting any such effect. A recent presentation in a meeting showed correlation between PD1.3 alleles and PDCD1 expression in peripheral blood cells.³² However, this study was done in samples from Iceland and Sweden, population showing the clearest association between the SNP and SLE susceptibility. This could reflect a specific pattern of linkage disequilibrium more than a specific effect of the SNP.

In conclusion, our results question the proposed regulatory function of the PD1.3 SNP and should be an incentive to search for other PDCD1 polymorphisms that could affect disease susceptibility. This is an important aim as PDCD1 seems to be involved in many different diseases and there are several pieces of evidence that indicate that genetic variation in PDCD1 is associated with susceptibility to some of them.

Acknowledgements

We thank Samuel Seoane-Ruzo (University of Santiago) for help with luciferase assays and Miguel A Iníguez-Peña (Centro de Biología Molecular, Madrid) for providing us with a working protocol for Jurkat cell transfection. This work was supported by two Grants, PI04/1615 and PI06/0620, from the Instituto de Salud Carlos III

(ISCIII, Spanish Health Ministry), which included funds from the FEDER programme of the European Union. MAS-G is the recipient of a pre-doctoral bursary from the Spanish Ministry of Education and Culture. IF-V is the recipient of an ISCIII pre-doctoral bursary. AG has been the recipient of a Research contract from the ISCIII.

References

- Prokunina L, Castillejo-Lopez C, Oberg F, Gunnarsson I, Berg L, Magnusson V et al. A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. *Nat Genet* 2002; **32**: 666–669.
- Velazquez-Cruz R, Orozco L, Espinosa-Rosales F, Carreno-Manjarrez R, Solis-Vallejo E, Lopez-Lara ND et al. Association of PDCD1 polymorphisms with childhood-onset systemic lupus erythematosus. *Eur J Hum Genet* 2007; **15**: 336–341.
- Thorburn CM, Prokunina-Olsson L, Sterba KA, Lum RF, Seldin MF, Alarcon-Riquelme ME et al. Association of PDCD1 genetic variation with risk and clinical manifestations of systemic lupus erythematosus in a multiethnic cohort. *Genes Immun* 2007; **8**: 279–287.
- Sigurðsson S, Nordmark G, Goring HH, Lindroos K, Wiman AC, Sturfelt G et al. Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. *Am J Hum Genet* 2005; **76**: 528–537.
- Ferreiros-Vidal I, D'Alfonso S, Papasteriades C, Skopouli FN, Marchini M, Scorza R et al. Bias in association studies of systemic lupus erythematosus susceptibility due to geographical variation in the frequency of a programmed cell death 1 polymorphism across Europe. *Genes Immun* 2007; **8**: 138–146.
- Nielsen C, Laustrop H, Voss A, Junker P, Husby S, Lillvang ST et al. A putative regulatory polymorphism in PD-1 is associated with nephropathy in a population-based cohort of systemic lupus erythematosus patients. *Lupus* 2004; **13**: 510–516.
- Sanghera DK, Manzi S, Bontempo F, Nestlerode C, Kamboh MI. Role of an intronic polymorphism in the PDCD1 gene with the risk of sporadic systemic lupus erythematosus and the occurrence of antiphospholipid antibodies. *Hum Genet* 2004; **115**: 393–398.
- Johansson M, Arlestig L, Moller B, Rantapaa-Dahlqvist S. Association of a PDCD1 polymorphism with renal manifestations in systemic lupus erythematosus. *Arthritis Rheum* 2005; **52**: 1665–1669.
- Ferreiros-Vidal I, Gomez-Reino JJ, Barros F, Carracedo A, Carreira P, Gonzalez-Escribano F et al. Association of PDCD1 with susceptibility to systemic lupus erythematosus: evidence of population-specific effects. *Arthritis Rheum* 2004; **50**: 2590–2597.
- Ferreiro-Neira I, Calaza M, Alonso-Perez E, Marchini M, Scorza R, Sebastiani GD et al. Opposed independent effects and epistasis in the complex association of IRF5 to SLE. *Genes Immun* 2007; **8**: 429–439.
- Nath SK, Kilpatrick J, Harley JB. Genetics of human systemic lupus erythematosus: the emerging picture. *Curr Opin Immunol* 2004; **16**: 794–800.
- Keir ME, Francisco LM, Sharpe AH. PD-1 and its ligands in T-cell immunity. *Curr Opin Immunol* 2007; **19**: 309–314.
- Gotsman I, Grabie N, Dacosta R, Sukhova G, Sharpe A, Lichtman AH et al. Proatherogenic immune responses are regulated by the PD-1/PD-L pathway in mice. *J Clin Invest* 2007; **117**: 2974–2982.
- Day CL, Kaufmann DE, Kiepiela P, Brown JA, Moodley ES, Reddy S et al. PD-1 expression on HIV-specific T cells is associated with T-cell exhaustion and disease progression. *Nature* 2006; **443**: 350–354.
- Urbani S, Amadei B, Tola D, Massari M, Schivazappa S, Missale G et al. PD-1 expression in acute hepatitis C virus (HCV) infection is associated with HCV-specific CD8 exhaustion. *J Virol* 2006; **80**: 11398–11403.
- Meyers S, Downing JR, Hiebert SW. Identification of AML-1 and the (8;21) translocation protein (AML-1/ETO) as sequence-specific DNA-binding proteins: the runt homology domain is required for DNA binding and protein-protein interactions. *Mol Cell Biol* 1993; **13**: 6336–6345.
- Otto F, Lubbert M, Stock M. Upstream and downstream targets of RUNX proteins. *J Cell Biochem* 2003; **89**: 9–18.
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM et al. The human genome browser at UCSC. *Genome Res* 2002; **12**: 996–1006.
- Benson G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 1999; **27**: 573–580.
- Hubbard TJ, Aken BL, Beal K, Ballester B, Caccamo M, Chen Y et al. Ensembl 2007. *Nucleic Acids Res* 2007; **35**: D610–D617.
- Miller W, Rosenbloom K, Hardison RC, Hou M, Taylor J, Raney B et al. 28-Way vertebrate alignment and conservation track in the UCSC Genome Browser. *Genome Res* 2007; **17**: 1797–1808.
- Puig-Kroger A, Lopez-Rodriguez C, Rellosio M, Sanchez-Elsner T, Nueda A, Munoz E et al. Polyomavirus enhancer-binding protein 2/core binding factor/acute myeloid leukemia factors contribute to the cell type-specific activity of the CD11a integrin gene promoter. *J Biol Chem* 2000; **275**: 28507–28512.
- Rahmann S, Muller T, Vingron M. On the power of profiles for transcription factor binding site detection. *Stat Appl Genet Mol Biol* 2003; **2**: Article7.
- Collas P, Dahl JA. Chop it, ChIP it, check it: the current status of chromatin immunoprecipitation. *Front Biosci* 2008; **13**: 929–943.
- Hollenhorst PC, Shah AA, Hopkins C, Graves BJ. Genome-wide analyses reveal properties of redundant and specific promoter occupancy within the ETS gene family. *Genes Dev* 2007; **21**: 1882–1894.
- Yochum GS, McWeeney S, Rajaraman V, Cleland R, Peters S, Goodman RH et al. Serial analysis of chromatin occupancy identifies beta-catenin target genes in colorectal carcinoma cells. *Proc Natl Acad Sci USA* 2007; **104**: 3324–3329.
- MacKenzie A, Quinn J. A serotonin transporter gene intron 2 polymorphic region, correlated with affective disorders, has allele-dependent differential enhancer-like properties in the mouse embryo. *Proc Natl Acad Sci USA* 1999; **96**: 15251–15255.
- Giambra V, Fruscalzo A, Giufre' M, Martinez-Labarga C, Favaro M, Rocchi M et al. Evolution of human IgH3'EC duplicated structures: both enhancers HSI,2 are polymorphic with variation of transcription factor's consensus sites. *Gene* 2005; **346**: 105–114.
- Bellizzi D, Rose G, Cavalcante P, Covello G, Dato S, De Rango F et al. A novel VNTR enhancer within the SIRT3 gene, a human homologue of SIR2, is associated with survival at oldest ages. *Genomics* 2005; **85**: 258–263.
- Thornell A, Hallberg B, Grundstrom T. Binding of SL3-3 enhancer factor 1 transcriptional activators to viral and chromosomal enhancer sequences. *J Virol* 1991; **65**: 42–50.
- Mori M, Yamada R, Kobayashi K, Kawaida R, Yamamoto K. Ethnic differences in allele frequency of autoimmune-disease-associated SNPs. *J Hum Genet* 2005; **50**: 264–266.
- Kristjansdottir H, Steinsson K, Gunnarsson I, Thorsteinsdottir T, Alarcon-Riquelme ME. Lower Expression Levels of the PD1 Immunoreceptor in SLE Patients and Lower Frequency of CD25+ CD4 T-cells. *Arthritis Rheum* 2007; **56**: S746 (Abs. 1952).

3. Influencia del componente genético en la variabilidad clínica del LES

Weak association of Systemic Lupus Erythematosus clinical features with susceptibility alleles: a case-only study

Marian Suarez-Gestal¹, Manuel Calaza¹, Myriam Liz¹, Josep Ordi-Ros², Eva Balada², Marc Bijl³, Cees G Kallenberg³, Chryssa Papasteriades⁴, Iris Kappou-Rigatou⁴, Patricia Carreira⁵, Fotini N. Skopouli⁶, Maria Mavromati⁶, Reinhold E. Schmidt⁷, Torsten Witte⁷, Emöke Endreffy⁸, Attila Kovacs⁹, Maurizio Marchini¹⁰, Raffaella Scorza¹⁰, Sergio Migliaresi¹¹, Gian Domenico Sebastiani¹², Maria Jose Santos¹³, Filipe Vinagre¹³, Ana Suarez¹⁴, Carmen Gutierrez¹⁴, Ignacio Rego¹⁵, Francisco J Blanco¹⁵, Nadia Barizzzone¹⁶, Sandra D'Alfonso¹⁶, Rudolf Pullmann Jr¹⁷, Rudolf Pullmann¹⁸, Sarka Ruzickova¹⁹, Ctibor Dostal²⁰, Juan J Gomez-Reino^{1,21} and Antonio Gonzalez¹

¹ Laboratorio de Investigacion 10 and Rheumatology Unit, Instituto de Investigacion Sanitaria - Hospital Clinico Universitario de Santiago, Santiago de Compostela, 15706 Spain. mariansg82@hotmail.com; mcalaza@gmail.com; myriam_liz@hotmail.com

² Internal Medicine, Research Laboratory in Autoimmune Diseases, Hospital Vall d'Hebron, 08035 Barcelona, Spain. ordi@vhebron.net; ebalada@ir.vhebron.net

³ Department of Rheumatology and Clinical Immunology, University Medical Center Groningen, 9713 Groningen, The Netherlands. m.bijl@int.umcg.nl; c.g.m.kallenberg@int.azg.nl

⁴ Department of Histocompatibility and Immunology, Evangelismos Hospital, 10676 Athens, Greece. cpapaste@otenet.gr; I.Kappou-Rigatou@excite.com

⁵ Rheumatology Department, Hospital 12 de Octubre, 28041 Madrid, Spain. carreira@h12o.es

⁶ Pathophysiology Department, Athens University Medical School, Athens 115 27, Greece. fskopouli@euroclinic.gr; ghatj@otenet.gr

⁷ Division of Clinical Immunology, Department of Internal Medicine of the Hannover Medical School, D-30625 Hannover, Germany. immunologie@mh-hannover.de; witte.torsten@mh-hannover.de

⁸ Paediatrics Department, Albert Szent-Györgyi Medical and Pharmaceutical Centre, University of Szeged, 6721 Hungary. ENDREFFY@pedia.szote.u-szeged.hu

⁹ Department of Rheumatology, Hospital of Hungarian Railways, H-5000 Szolnok, Hungary. attila.kovacs.dr@freemail.hu

¹⁰ Clinical Immunology, University of Milan and Fondazione IRCCS Ospedale Maggiore Policlinico, Mangiagalli e Regina Elena, 20122 Milan, Italy.

maurizio.marchini@unimi.it; raffaella.scorza@unimi.it

¹¹ Rheumatology Unit, Second University of Naples, 81100 Naples, Italy.

sergio.migliaresi@unina2.it

¹² UOC Reumatologia, Azienda Ospedaliera San Camillo-Forlanini, Roma, Italy.

gsebastiani@scamilloforlanini.rm.it

¹³ Rheumatology Department, Hospital Garcia de Orta and Rheumatology Research Unit, Instituto Medicina Molecular, Lisboa, Portugal. MJPS@netvisao.pt;

filipevinagre@netcabo.pt

¹⁴ Department of Functional Biology, Hospital Universitario Central de Asturias, Universidad de Oviedo, Oviedo 33006, Spain. anasua@uniovi.es; cgut@hca.es

¹⁵ INIBIC-CH Universitario A Coruña, 15006 A Coruña, Spain. irego@canalejo.org;

Francisco_Blanco@canalejo.org

¹⁶ Dept Medical Sciences and IRCAD , Eastern Piedmont University, 28100 Novara, Italy. barizzo@med.unipmn.it; dalfonso@med.unipmn.it

¹⁷ GBMC Hospital, Department of Medicine, Baltimore, MD 21209, USA.

rudolf.pullmann@gmail.com

¹⁸ Institute of Clinical Biochemistry, Martin Faculty Hospital, Jessenius Medical Faculty, Kollárova 2, 036 59 Martin, Slovakia. pullmann@mfn.sk

¹⁹ Institute of Biotechnology, Academy of Sciences of the Czech Republic, Prague, Czech Republic. sarka.ruzickova@img.cas.cz

²⁰ Molecular Biology and Immunogenetics Dept, Institute of Rheumatology, 128 50 Prague 2, Czech Republic. dost@revma.cz

²¹ Department of Medicine, University of Santiago de Compostela, Santiago de Compostela, 15706 Spain. juan.gomez-reino.carnota@sergas.es

Corresponding author:

Antonio Gonzalez

Laboratorio de Investigacion 10

Hospital Clinico Universitario de Santiago

Travesia de Choupana sn.

15706-Santiago de Compostela

Spain

Tfn: 34 981 950 903

Fax: 34 981 951 068

Antonio.Gonzalez.Martinez-Pedrayo@sergas.es

anlugon@hotmail.com

ABSTRACT

Introduction: To analyze if nine newly identified SLE susceptibility loci were influencing the presence of some of the best defined SLE clinical features.

Methods: We compared the genotypes of a representative SNP in each of nine SLE-associated loci, in *ITGAM*, *STAT4*, *C8orf13-BLK*, *IRAK1-MECP2*, *BANK1*, *TYK2*, *KIAA1542*, *PXK* and *1q25.1*, between patients with and without each of 11 clinical features: ten of the ACR classification criteria (except ANAs) and age of disease onset. Data of 1561 patients were used.

Results: The most consistent and significant result was the association of the rare allele of *rs13277113* in the *C8orf13-BLK* locus with increased risk of nephritis (O.R. = 1.35, P = 0.00026). The reported association of *STAT4* with early age of disease onset was replicated. Not so the other described associations of this locus, but joint analysis of a previous report with our data lent further support to association of *STAT4* with oral ulcers and immunologic disorder. Combined analysis of *ITGAM* data from a recent report and our study showed more doubtful association of this locus with immunologic disorder, nephritis and discoid rash, although our study, separately, did not replicate these associations. Some other results were suggestive of the presence of additional phenotype-genotype associations, but they were not significant after correction for the number of tests.

Conclusions: Our results together with previous studies indicate that clinical variability of SLE is influenced by the SLE susceptibility alleles. However, the effects are weak and difficult to reproduce indicating the need of phenotype-specific genetic studies.

INTRODUCTION

Systemic Lupus Erythematosus (SLE; OMIM 152700) is a complex disease with heterogeneous clinical manifestations and a genetic component that begins to be understood [1]. Clinical heterogeneity complicates its diagnosis, treatment and research [2]. We do not have yet clear clues about the causes of this variability, but likely contributing elements are stochastic factors and differences in the environmental exposures or in the genetic constitution. Understanding the relationships between these factors and the SLE features will help elucidate disease mechanisms and could provide the basis for a classification of patients in more homogeneous subgroups.

Many associations between genetic polymorphisms and the presence of specific SLE clinical features have been reported, but they have not yet reached a high degree of confidence [3-6]. This is similar to what had happened in relation to SLE susceptibility until 2008: although there were many reports of association there were few genetic factors that had been consistently demonstrated. The advent of Genome Wide Association (GWA) studies in 2008 and other large studies, with their increased sample sizes and stricter standards, changed dramatically this panorama. Now, there are more than ten new loci that are consistently associated with SLE susceptibility [7-8]. These loci are prime candidates for having a predisposing effect for specific SLE clinical features. The first analyses seem to indicate that this is the case for some of them [9-11]. Herein, we have analyzed the association of nine of these consistent SLE susceptibility loci (*ITGAM*, *STAT4*, *C8orf13-BLK*, *IRAK1-MECP2*, *BANK1*, *TYK2*, *KIAA1542*, *PXK* and *lq25.1*) with some clinical features in more than 1500 patients. Some significant associations were found and some of the previously described received some additional support, but the observed effects were weak and will need further confirmation.

MATERIAL AND METHODS

Sample and data collection: Recruiting of DNA samples from SLE patients has been already described [12]. Briefly, 16 recruiting centres from 9 European countries were asked for about 100 SLE patients and controls of uniform Caucasian ethnicity. A total of 1561 patients were obtained (18 patients included in a previous study[12] were excluded after reviewing their clinical data). All SLE patients met the revised American College of Rheumatology classification criteria [13], and gave written informed consent. Sample collection was approved by the respective ethical committees. Clinical characteristics of the patients were obtained at the same time. They included the ACR classification criteria, gender, age of disease onset, and time of follow-up.

Genotypes: Genotypes for the SLE-associated SNPs in *ITGAM*, *STAT4*, *C8orf13-BLK*, *IRAK1-MECP2*, *BANK1*, *TYK2*, *KIAA1542*, *PXK* and *1q25.1* loci have been already described [12].

Association analysis: We compared SNP allele frequencies between cases positive and negative for each of the ACR classification criteria (except presence of ANA given that they were almost constant) and age of disease onset by logistic regression following a genetic additive model with codes 2 for aa genotypes, 1 for aA genotypes and 0 for the AA genotypes. Patient gender was included as covariate. Age of disease onset was included in these analysis after separating patients in two groups $< \text{or} \geq 30$ years old at disease onset. Results at this stage with $P < 0.05$ were analyzed in patients stratified by gender. Bonferroni correction for multiple tests was done considering the number of analyzed clinical features. Statistical analyses were done in a customized version of the Statistica 7.0 program (StatSoft, Tulsa, OK, USA). Power to detect association was estimated with the “Power and sample size” software [14]. Combination of the current

results with previous studies was done with fixed-effect and with random-effect metaanalyses as implemented in the R project (<http://www.r-project.org/>).

RESULTS

Characteristics of the patient collection

The genotypes of strongly associated SNPs in 9 confirmed SLE susceptibility loci were available for 1561 European patients with SLE and for 1728 healthy controls (Table 1). Information about ACR classification criteria, gender, age of disease onset, and time of follow-up was also available for most of these patients (Table 2).

Case-only analysis of the association with SLE clinical features

Seven associations between genotypes and clinical features were detected below $P = 0.05$ in the comparison between positive and negative patients (Table 3). Only two of them were below $P = 0.01$. The two involved nephritis that was associated with the SLE risk allele at the *C8orf13-BLK* locus and with the *IRAK1-MECP2* risk allele. The other five associations were just below the 0.05 threshold. One of the weak associations replicates a previous finding: the association of the risk allele of the *STAT4* SNP with younger disease onset [9]. However, we did not replicate the described associations of the *STAT4* SNP with nephritis, immunologic disorder or oral ulcers (although there was a trend in the same direction for the latter) [9-10]. Other weak association was observed between the *ITGAM* risk allele and a younger age of disease onset, but we did not replicate other reported associations with this SLE risk allele [11]. We also found weak association of renal disorder with the risk allele in *BANK1* and of oral ulcers with this same allele and the risk *PXX* allele (Table 3). However, only the association of nephritis with the *C8orf13-BLK* SNP remained significant after correcting by the number of tests.

Analyses of patients stratified by gender was done only for the SNPs showing $P < 0.05$ in global analyses (Table 4). The smallest stratum included the 151 men with SLE in our collection, what allowed for a very limited statistical power. This limitation was considered for interpretation by focusing in effect sizes more than in P values. In contrast, the largest stratum was made of 1393 women with SLE, and therefore similar to the global study sample. The effect size of five of the associations was similar or stronger in males than in females and concordant in direction. These associations included the two stronger, renal disorder with the risk alleles of *C8orf13-BLK* and of *IRAK1-MECP2*. The other two significant associations in the global analysis, renal disorder with *BANK1* and oral ulcers with *PXK*, were restricted to women in the stratified analysis (Table 3).

Joint analysis with previous data

We combined our results with the available in previous reports. There were few articles and only in three of them data were reported as to permit case-only analyses, comparing patients with and without the clinical feature. Each of these three articles provided data on one of the loci included in our analyses, *C8orf13-BLK*, *STAT4* and *ITGAM* [9, 11, 15]. There was an additional study analyzing SLE clinical features in relation with the *ITGAM* locus but with a different SNP [15]. The other published studies did not provide enough information to combine effect sizes, as a second one about *STAT4* [10], or the studies on *BANK1* [16], and *TYK2* [17]; or there was not any report, as for the remaining loci.

The clearest association in our study, that of *rs13277113* in *C8orf13-BLK* with renal disorder, was opposed to the results of a GWAS that included clinical data on 781 patients [15]. The marked contrast between the two studies leaves this association as uncertain at this time (Table 5). On the contrary, the results for the *STAT4* locus were

quite similar between the two available studies. The associations between young age of disease onset, oral ulcers or immunologic disorder and *rs7574865* were concordant and, consequently, they became reinforced in the combined analysis with about 2700 patients (Table 5). Only the association that was previously observed between *rs7574865* and renal disorder did not receive any additional support, and it became non-significant ($P = 0.14$) in the combined analysis. Finally, less conclusive results were obtained with the joint analysis on the *ITGAM* locus. The direction of change for the three analyzed criteria was the same in the two studies, but results from Kim-Howard et al. showed stronger differences than ours. The differences in effects for renal disorder and discoid rash were such as to cause conflicting results between the fixed-effects and random-effects metaanalyses. There were significant associations with the fixed-effects analysis and lack of them with the more conservative random-effects approach. Immunologic disorder was associated with $P = 0.04$ with both approaches.

DISCUSSION

There are several obstacles for the identification of loci preferentially associated with particular SLE phenotypes that make their investigation more difficult than the study of disease susceptibility. Firstly, the statistical power of the available collections becomes greatly reduced in subgroup analyses. Lack of power will lead to false negative results. In addition, the performance of many comparisons increases the risk of false positives and the subsequent lack of reproducibility. A solution to this problem will be to require stringent significant thresholds for claiming association, but this measure will lead to even more false negative results. Another limitation is the variability of the clinical data, because sample collections show marked differences in the prevalence of the different SLE manifestations [9, 15, 18-23]. These differences have many possible

causes, but if they are due to biases unrelated with the disease process they will be a source of spurious associations [24]. Consideration of these obstacles and the recommendations for subgroup analysis in clinical trials [25-27], have determined the design of our study. We have followed a fixed sequence approach, which requires a significant difference in the more general group before starting analysis in a subgroup; we have also corrected for the number of tests done, took into consideration previous reports concerning the same comparisons, and take into account the power of the analyses for interpretation.

According to the fixed sequence approach, we selected nine loci that are convincingly associated to SLE susceptibility and that have shown association in our collection of samples [7-8, 12]. This procedure does not exclude that some gene variants could predispose to a particular SLE clinical phenotype without increasing appreciably susceptibility to SLE overall. This has been already reported for the *FCGR3A V158F* SNP that is associated with lupus nephritis, but not convincingly with SLE susceptibility in patients without nephritis [5].

The clearest associations in our study were with renal disorder. Association between lupus nephritis and the SLE risk allele of the *rs13277113* SNP in the *C8orf13-BLK* locus remained significant after correcting for the number of tests, and was consistently observed in women and men. The same SNP has already been analyzed for association with SLE classification criteria in a previous GWAS [15], but no association with renal disorder was found in any of the two series of patients analyzed in that study. In addition, the direction of change in one of the series, which was reported in more detail, was opposed to the observed by us (Table 5). Prevalence of renal disorder in that study was lower (28.0 %) than in our collection of samples (41.0 %) and this difference ($P = 2.0 \times 10^{-12}$) could contribute to the discordant results. The *C8orf13-BLK* locus is

reproducibly associated with SLE susceptibility [12, 15, 28], and probably with RA [29-30], and Systemic Sclerosis [31]. The SNP that we have analyzed is the most studied and seems to be the most associated with SLE susceptibility. It correlates with changes in expression of the two flanking genes, *C8orf13* and *BLK*, with an increase of *C8orf13* and a decrease of *BLK* [15]. *BLK* is the best understood of the two, and it is a good candidate because it codes for a B-cell specific tyrosine kinase that is involved in maturation of B cells and in BCR signalling [32-33]. A second association of lupus nephritis in our study was less clear. The SLE risk allele of *rs17435* in the *IRAK1-MECP2* locus was associated with increased risk of nephritis both in women and men. However, the association did not remain significant after correction by the number of tests done. Therefore, we consider these two associations of renal disorder as unproven until other studies provide additional evidence.

The *rs7574865* SNP in *STAT4* has been associated with a severe SLE phenotype defined by nephritis, age at diagnosis < 30 years old, immunologic disorder (and, specifically, double-stranded DNA autoantibodies) and absence of oral ulcers [9]. The association with immunologic disorder (and anti-dsDNA) and a modest increase in nephritis prevalence was also reported in other study with 695 patients [10]. However, no effect on age of disease onset or oral ulcers was detected in this second study. Our results showed only a weak association with younger disease onset that was consistent in women and men. We also found a trend to decreased prevalence of oral ulcers, but it was only observed in women (data not shown). No association with nephritis or immunologic disorder was present in our data. Overall, the results confirm that *rs7574865* SNP in the *STAT4* locus is associated with a particular SLE phenotype whose details are still incompletely defined, but that very likely include immunologic disorder (same trend in the three studies), specially the presence of anti-dsDNA

antibodies, which was a clear association in the two previous studies and was not available to us; a younger age of disease onset, for which there was weak association in two studies and a non-significant trend in the third [10]. It is also probable that this phenotype will include also renal disorder and lack of oral ulcers because two of the three studies were concordant for these two associations.

We also found association of *rs1143679*, which is a likely causal non-synonymous SNP in *ITGAM*, with a younger age of disease onset. This characteristic of SLE has not been analyzed in previous studies [11, 15], and the association did not remain significant after correction by the number of tests, which make this association of uncertain relevance. Something similar happens with other recently reported clinical associations with this locus [11]. Joint analysis with our data show only weak association with immunologic disorder ($P = 0.04$) and inconclusive results regarding renal disorder and discoid rash due to the differences in effect size between the two studies. In addition, the other available report shows no association with any of these three clinical features in any of the two series of patients considered [15], albeit the analyzed SNP in that report, *rs11574637*, is only partially correlated with *rs1143679*.

Of the remaining loci studied here, only *TYK2* and *BANK1* have previously been examined for association with SLE clinical features [16-17]. No specific association was found in the previous studies and we did not find any in ours, in spite of relative good power. In the interpretation of these negative results, it is important to consider the power of the each comparison. All the SNPs analyzed here have similar power because this is determined by the minor allele frequencies that are similar. However, the clinical features show a gradation of power in function of the rarity of the phenotype with lowest frequency (Table 2). Larger power is available for clinical features showing a prevalence approaching 50%, like age of disease onset < 30, photosensitivity or malar

rash, whereas a lower power corresponds to clinical features with low prevalence, like neurologic disorder, or with very high prevalence, like arthritis or female gender. None of the latter was associated in our study or has been reported as associated in previous studies.

Our results and previous reports on the same loci [9-10, 15-17] and in other SLE susceptibility loci [17, 34-41] do not show a clear and consistent association with specific clinical features or subphenotypes. Association of the *STAT4* locus with some SLE features is at present the only significant exception to this trend. This conclusion is not contradicted by some studies that have found differences in SLE susceptibility loci between patients with a clinical feature and healthy controls, but not differences between patients with and without the clinical manifestation [16, 39, 41]. Other confirmed SLE susceptibility loci that have not shown association with specific clinical manifestations include *IRF5* [17, 36-38], *PTPN22* [34-35, 39-40] and *TNFAIP3* [41]. The same can be affirmed for the *MHC*, except that in this case, convincing associations with particular autoantibodies have been reported [42-45]. The available evidence does not exclude, however, that future studies will find modest or weak associations that have escaped detection due to lack of power of the current SLE patient collections.

In retrospect, a weak phenotype specificity of the SLE susceptibility loci is coherent with the fact that they have been identified in studies involving all types of patients with SLE. It is likely that large phenotype-specific GWAs will uncover genetic variants with a more clear effect on phenotypes, even if they have less impact in SLE overall. This type of studies should be actively promoted in the hope that will allow identifying more uniform subgroups of patients with SLE, which could be very useful to improve disease management. Also, it is possible that complex interactions between loci or with environmental exposures play a more significant role in the phenotypic

heterogeneity of SLE. In any case, a contribution of genetic variation to SLE phenotypes is supported by ethnic [46-47], ancestry [48] and family studies [49-51], and it should be expected that future research will be able to identify the phenotype-specific loci.

CONCLUSIONS

The paucity of sound associations with specific clinical features in our study, together with other reports on the same loci [9-10, 15-17], or on other confirmed SLE susceptibility loci [17, 34-41] shows that clinical variability of SLE does not correlate in a consistent and clear way with SLE susceptibility alleles. The best of the already known associations is with the *STAT4* locus, and it is still incompletely defined. Other associations that are only tentative at present will be confirmed in new studies, but it is likely that only phenotype-specific GWAS will lead to the discovery of a significant fraction of their genetic component.

ABBREVIATIONS

SLE: Systemic Lupus Erythematosus

GWA: Genome Wide Association study

anti-dsDNA: Double-stranded DNA autoantibodies

MAF: Minor allele frequency

CI: Confidence Interval

OR: Odds ratio

COMPETING INTEREST

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

MS-G participated in design of the study, genotyped the samples, participated in the interpretation of the results and in writing the manuscript. MC participated in the statistical analysis and in the interpretation of results. ML, JO-R, EB, MB, C-GK, CP, IK-R, PC, F-NS, MM, R-ES, TW, EE, AK, MM, RS, SM, GD-S, M-JS, FV, AS, CG, IR, F-JB, NB, SD, RP Jr, RP, SR, CD, and JJG-R participated in the acquisition of clinical data, collection of samples and in the analysis and interpretation of results. AG participated in the design of the study, coordinated acquisition of clinical data and collection of samples, supervised genotyping, statistical analysis, interpretation of results and writing of the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

MS-G is the recipient of a FPU pre-doctoral bursary of the Spanish Ministry of Education. The present work was supported by Fondo de Investigacion Sanitaria of the Instituto de Salud Carlos III (Spain), grants 06/0620, 08/0744 and RD08/0075/0019 that are partially financed by the Fondo Europeo de Desarrollo Regional program of the European Union, by grants from the Xunta de Galicia, and by BMBF KN Rheuma grant C2.12 (to TW)

REFERENCES

1. Wallace DJ, Hahn B, Dubois EL: **Dubois' lupus erythematosus**, 7th edn. Philadelphia: Lippincott Williams & Wilkin; 2007.

2. Eisenberg R: **Why can't we find a new treatment for SLE?** *J Autoimmun* 2009, **32**(3-4):223-230.
3. Tsao BP: **Update on human systemic lupus erythematosus genetics.** *Curr Opin Rheumatol* 2004, **16**(5):513-521.
4. Lauwerys BR, Wakeland EK: **Genetics of lupus nephritis.** *Lupus* 2005, **14**(1):2-12.
5. Karassa FB, Trikalinos TA, Ioannidis JP: **The Fc gamma RIIIA-F158 allele is a risk factor for the development of lupus nephritis: a meta-analysis.** *Kidney Int* 2003, **63**(4):1475-1482.
6. Sestak AL, Nath SK, Sawalha AH, Harley JB: **Current status of lupus genetics.** *Arthritis Res Ther* 2007, **9**(3):210.
7. Moser KL, Kelly JA, Lessard CJ, Harley JB: **Recent insights into the genetic basis of systemic lupus erythematosus.** *Genes Immun* 2009, **10**(5):373-379.
8. Rhodes B, Vyse TJ: **The genetics of SLE: an update in the light of genome-wide association studies.** *Rheumatology (Oxford)* 2008, **47**(11):1603-1611.
9. Taylor KE, Remmers EF, Lee AT, Ortmann WA, Plenge RM, Tian C, Chung SA, Nititham J, Hom G, Kao AH, Demirci FY, Kamboh MI, Petri M, Manzi S, Kastner DL, Seldin MF, Gregersen PK, Behrens TW, Criswell LA: **Specificity of the STAT4 genetic association for severe disease manifestations of systemic lupus erythematosus.** *PLoS Genet* 2008, **4**(5):e1000084.
10. Sigurdsson S, Nordmark G, Garnier S, Grundberg E, Kwan T, Nilsson O, Eloranta ML, Gunnarsson I, Svenungsson E, Sturfelt G, Bengtsson AA, Jonsen A, Truedsson L, Rantapaa-Dahlqvist S, Eriksson C, Alm G, Goring HH, Pastinen T, Syvanen AC, Ronnblom L: **A risk haplotype of STAT4 for systemic lupus erythematosus is over-expressed, correlates with anti-**

- dsDNA and shows additive effects with two risk alleles of IRF5.** *Hum Mol Genet* 2008, **17**(18):2868-2876.
11. Kim-Howard X, Maiti AK, Anaya JM, Bruner GR, Brown E, Merrill JT, Edberg JC, Petri MA, Reveille JD, Ramsey-Goldman R, Alarcon GS, Vyse TJ, Gilkeson G, Kimberly RP, James JA, Guthridge JM, Harley JB, Nath SK: **ITGAM coding variant (rs1143679) influences the risk of renal disease, discoid rash, and immunologic manifestations in lupus patients with European ancestry.** *Ann Rheum Dis* 2009.
 12. Suarez-Gestal M, Calaza M, Endreffy E, Pullmann R, Ordi-Ros J, Domenico Sebastiani G, Ruzickova S, Jose Santos M, Papasteriades C, Marchini M, Skopouli FN, Suarez A, Blanco FJ, D'Alfonso S, Bijl M, Carreira P, Witte T, Migliaresi S, Gomez-Reino JJ, Gonzalez A: **Replication of recently identified systemic lupus erythematosus genetic associations: a case-control study.** *Arthritis Res Ther* 2009, **11**(3):R69.
 13. Hochberg MC: **Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus.** *Arthritis Rheum* 1997, **40**(9):1725.
 14. Dupont WD, Plummer WD, Jr.: **Power and sample size calculations. A review and computer program.** *Control Clin Trials* 1990, **11**(2):116-128.
 15. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, Lee AT, Chung SA, Ferreira RC, Pant PV, Ballinger DG, Kosoy R, Demirci FY, Kamboh MI, Kao AH, Tian C, Gunnarsson I, Bengtsson AA, Rantapaa-Dahlqvist S, Petri M, Manzi S, Seldin MF, Ronnblom L, Syvanen AC, Criswell LA, Gregersen PK, Behrens TW: **Association of systemic lupus**

- erythematosus with C8orf13-BLK and ITGAM-ITGAX.** *N Engl J Med* 2008, **358**(9):900-909.
16. Guo L, Deshmukh H, Lu R, Vidal GS, Kelly JA, Kaufman KM, Dominguez N, Klein W, Kim-Howard X, Bruner GR, Scofield RH, Moser KL, Gaffney PM, Dozmorov IM, Gilkeson GS, Wakeland EK, Li QZ, Langefeld CD, Marion MC, Williams AH, Divers J, Alarcon GS, Brown EE, Kimberly RP, Edberg JC, Ramsey-Goldman R, Reveille JD, McGwin G, Jr., Vila LM, Petri MA *et al*: **Replication of the BANK1 genetic association with systemic lupus erythematosus in a European-derived population.** *Genes Immun* 2009, **10**(5):531-538.
17. Sigurdsson S, Nordmark G, Goring HH, Lindroos K, Wiman AC, Sturfelt G, Jonsen A, Rantapaa-Dahlqvist S, Moller B, Kere J, Koskenmies S, Widen E, Eloranta ML, Julkunen H, Kristjansdottir H, Steinsson K, Alm G, Ronnblom L, Syvanen AC: **Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus.** *Am J Hum Genet* 2005, **76**(3):528-537.
18. Cervera R, Abarca-Costalago M, Abramovicz D, Allegri F, Annunziata P, Aydintug AO, Bacarelli MR, Bellisai F, Bernardino I, Biernat-Kaluza E, Blockmans D, Boki K, Bracci L, Campanella V, Camps MT, Carcassi C, Cattaneo R, Cauli A, Chwalinska-Sadowska H, Contu L, Cosyns JP, Danieli MG, D DC, Depresseux G, Direskeneli H, Domenech I, Espinosa G, Fernandez-Nebro A, Ferrara GB, Font J *et al*: **Systemic lupus erythematosus in Europe at the change of the millennium: lessons from the "Euro-Lupus Project".** *Autoimmun Rev* 2006, **5**(3):180-186.

19. Pons-Estel BA, Catoggio LJ, Cardiel MH, Soriano ER, Gentiletti S, Villa AR, Abadi I, Caeiro F, Alvarellos A, Alarcon-Segovia D: **The GLADEL multinational Latin American prospective inception cohort of 1,214 patients with systemic lupus erythematosus: ethnic and disease heterogeneity among "Hispanics"**. *Medicine (Baltimore)* 2004, **83**(1):1-17.
20. Alarcon GS, McGwin G, Jr., Petri M, Reveille JD, Ramsey-Goldman R, Kimberly RP: **Baseline characteristics of a multiethnic lupus cohort: PROFILE**. *Lupus* 2002, **11**(2):95-101.
21. Petri M: **The effect of race on the presentation and course of SLE in the United States**. *Arthritis and Rheumatism* 1997, **40**(9):785-785.
22. Wang F, Wang CL, Tan CT, Manivasagar M: **Systemic lupus erythematosus in Malaysia: a study of 539 patients and comparison of prevalence and disease expression in different racial and gender groups**. *Lupus* 1997, **6**(3):248-253.
23. Al Arfaj AS, Khalil N: **Clinical and immunological manifestations in 624 SLE patients in Saudi Arabia**. *Lupus* 2009, **18**(5):465-473.
24. Hulley SB: **Designing clinical research**, 3rd edn. Philadelphia, PA: Lippincott Williams & Wilkins; 2007.
25. Guillemin F: **Primer: the fallacy of subgroup analysis**. *Nat Clin Pract Rheumatol* 2007, **3**(7):407-413.
26. Brookes ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey Smith G: **Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives**. *Health Technol Assess* 2001, **5**(33):1-56.
27. Moyé LA: **Multiple analyses in clinical trials: fundamentals for investigators**. New York: Springer; 2003.

28. Harley JB, Alarcon-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, Moser KL, Tsao BP, Vyse TJ, Langefeld CD, Nath SK, Guthridge JM, Cobb BL, Mirel DB, Marion MC, Williams AH, Divers J, Wang W, Frank SG, Namjou B, Gabriel SB, Lee AT, Gregersen PK, Behrens TW, Taylor KE, Fernando M, Zidovetzki R, Gaffney PM, Edberg JC, Rioux JD, Ojwang JO *et al*: **Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci.** *Nat Genet* 2008, **40**(2):204-210.
29. Suarez-Gestal M, Calaza M, Dieguez-Gonzalez R, Perez-Pampin E, Pablos JL, Navarro F, Narvaez J, Marengo JL, Herrero-Beaumont G, Fernandez-Gutierrez B, Lamas JR, de la Serna AR, Ortiz AM, Carreno L, Canete JD, Caliz R, Blanco FJ, Balsa A, Gomez-Reino JJ, Gonzalez A: **Rheumatoid arthritis does not share most of the newly identified systemic lupus erythematosus genetic factors.** *Arthritis Rheum* 2009, **60**(9):2558-2564.
30. Gregersen PK, Amos CI, Lee AT, Lu Y, Remmers EF, Kastner DL, Seldin MF, Criswell LA, Plenge RM, Holers VM, Mikuls TR, Sokka T, Moreland LW, Bridges SL, Jr., Xie G, Begovich AB, Siminovitch KA: **REL, encoding a member of the NF-kappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis.** *Nat Genet* 2009, **41**(7):820-823.
31. Gourh P, Agarwal SK, Martin E, Divecha D, Rueda B, Bunting H, Assassi S, Paz G, Shete S, McNearney T, Draeger H, Reveille JD, Radstake TR, Simeon CP, Rodriguez L, Vicente E, Gonzalez-Gay MA, Mayes MD, Tan FK, Martin J, Arnett FC: **Association of the C8orf13-BLK region with systemic sclerosis in North-American and European populations.** *J Autoimmun* 2009.

32. Dymecki SM, Niederhuber JE, Desiderio SV: **Specific expression of a tyrosine kinase gene, blk, in B lymphoid cells.** *Science* 1990, **247**(4940):332-336.
33. Saijo K, Schmedt C, Su IH, Karasuyama H, Lowell CA, Reth M, Adachi T, Patke A, Santana A, Tarakhovsky A: **Essential role of Src-family protein tyrosine kinases in NF-kappaB activation during B cell development.** *Nat Immunol* 2003, **4**(3):274-279.
34. Orozco G, Sanchez E, Gonzalez-Gay MA, Lopez-Nevot MA, Torres B, Caliz R, Ortego-Centeno N, Jimenez-Alonso J, Pascual-Salcedo D, Balsa A, de Pablo R, Nunez-Roldan A, Gonzalez-Escribano MF, Martin J: **Association of a functional single-nucleotide polymorphism of PTPN22, encoding lymphoid protein phosphatase, with rheumatoid arthritis and systemic lupus erythematosus.** *Arthritis Rheum* 2005, **52**(1):219-224.
35. Kyogoku C, Langefeld CD, Ortmann WA, Lee A, Selby S, Carlton VE, Chang M, Ramos P, Baechler EC, Batliwalla FM, Novitzke J, Williams AH, Gillett C, Rodine P, Graham RR, Ardlie KG, Gaffney PM, Moser KL, Petri M, Begovich AB, Gregersen PK, Behrens TW: **Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE.** *Am J Hum Genet* 2004, **75**(3):504-507.
36. Cunninghame Graham DS, Manku H, Wagner S, Reid J, Timms K, Gutin A, Lanchbury JS, Vyse TJ: **Association of IRF5 in UK SLE families identifies a variant involved in polyadenylation.** *Hum Mol Genet* 2007, **16**(6):579-591.
37. Sigurdsson S, Goring HH, Kristjansdottir G, Milani L, Nordmark G, Sandling JK, Eloranta ML, Feng D, Sangster-Guity N, Gunnarsson I, Svenungsson E, Sturfelt G, Jonsen A, Truedsson L, Barnes BJ, Alm G, Ronnblom L, Syvanen AC: **Comprehensive evaluation of the genetic variants of interferon**

- regulatory factor 5 (IRF5) reveals a novel 5 bp length polymorphism as strong risk factor for systemic lupus erythematosus.** *Hum Mol Genet* 2008, **17**(6):872-881.
38. Ferreiro-Neira I, Calaza M, Alonso-Perez E, Marchini M, Scorza R, Sebastiani GD, Blanco FJ, Rego I, Pullmann R, Jr., Pullmann R, Kallenberg CG, Bijl M, Skopouli FN, Mavromati M, Migliaresi S, Barizzzone N, Ruzickova S, Dostal C, Schmidt RE, Witte T, Papasteriades C, Kappou-Rigatou I, Endreffy E, Kovacs A, Ordi-Ros J, Balada E, Carreira P, Gomez-Reino JJ, Gonzalez A: **Opposed independent effects and epistasis in the complex association of IRF5 to SLE.** *Genes Immun* 2007, **8**(5):429-438.
39. Reddy MV, Johansson M, Sturfelt G, Jonsen A, Gunnarsson I, Svenungsson E, Rantapaa-Dahlqvist S, Alarcon-Riquelme ME: **The R620W C/T polymorphism of the gene PTPN22 is associated with SLE independently of the association of PDCD1.** *Genes Immun* 2005, **6**(8):658-662.
40. Balada E, Villarreal-Tolchinsky J, Ordi-Ros J, Labrador M, Serrano-Acedo S, Martinez-Lostao L, Vilardell-Tarres M: **Multiplex family-based study in systemic lupus erythematosus: association between the R620W polymorphism of PTPN22 and the FcgammaRIIa (CD32A) R131 allele.** *Tissue Antigens* 2006, **68**(5):432-438.
41. Bates JS, Lessard CJ, Leon JM, Nguyen T, Battiest LJ, Rodgers J, Kaufman KM, James JA, Gilkeson GS, Kelly JA, Humphrey MB, Harley JB, Gray-McGuire C, Moser KL, Gaffney PM: **Meta-analysis and imputation identifies a 109 kb risk haplotype spanning TNFAIP3 associated with lupus nephritis and hematologic manifestations.** *Genes Immun* 2009, **10**(5):470-477.

42. Galeazzi M, Sebastiani GD, Morozzi G, Carcassi C, Ferrara GB, Scorza R, Cervera R, de Ramon Garrido E, Fernandez-Nebro A, Houssiau F, Jedryka-Goral A, Passiu G, Papasteriades C, Piette JC, Smolen J, Porciello G, Marcolongo R: **HLA class II DNA typing in a large series of European patients with systemic lupus erythematosus: correlations with clinical and autoantibody subsets.** *Medicine (Baltimore)* 2002, **81**(3):169-178.
43. Galeazzi M, Sebastiani GD, Tincani A, Piette JC, Allegri F, Morozzi G, Bellisai F, Scorza R, Ferrara GB, Carcassi C, Font J, Passiu G, Smolen J, Papasteriades C, Houssiau F, Nebro AF, Ramon Garrido ED, Jedryka-Goral A, Marcolongo R: **HLA class II alleles associations of anticardiolipin and anti-beta2GPI antibodies in a large series of European patients with systemic lupus erythematosus.** *Lupus* 2000, **9**(1):47-55.
44. Vasconcelos C, Carvalho C, Leal B, Pereira C, Bettencourt A, Costa PP, Marinho A, Barbosa P, Almeida I, Farinha F, Mendonca T, Correia JA, Mendonca D, Martins B: **HLA in Portuguese systemic lupus erythematosus patients and their relation to clinical features.** *Ann N Y Acad Sci* 2009, **1173**:575-580.
45. Fernando MM, Stevens CR, Walsh EC, De Jager PL, Goyette P, Plenge RM, Vyse TJ, Rioux JD: **Defining the role of the MHC in autoimmunity: a review and pooled analysis.** *PLoS Genet* 2008, **4**(4):e1000024.
46. Uribe AG, McGwin G, Jr., Reveille JD, Alarcon GS: **What have we learned from a 10-year experience with the LUMINA (Lupus in Minorities; Nature vs. nurture) cohort? Where are we heading?** *Autoimmun Rev* 2004, **3**(4):321-329.

47. Peschken CA, Katz SJ, Silverman E, Pope JE, Fortin PR, Pineau C, Smith CD, Arbillaga HO, Gladman DD, Urowitz M, Zummer M, Clarke A, Bernatsky S, Hudson M: **The 1000 Canadian faces of lupus: determinants of disease outcome in a large multiethnic cohort.** *J Rheumatol* 2009, **36**(6):1200-1208.
48. Chung SA, Tian C, Taylor KE, Lee AT, Ortmann WA, Hom G, Graham RR, Nititham J, Kelly JA, Morrissey J, Wu H, Yin H, Alarcon-Riquelme ME, Tsao BP, Harley JB, Gaffney PM, Moser KL, Manzi S, Petri M, Gregersen PK, Langefeld CD, Behrens TW, Seldin MF, Criswell LA: **European population substructure is associated with mucocutaneous manifestations and autoantibody production in systemic lupus erythematosus.** *Arthritis Rheum* 2009, **60**(8):2448-2456.
49. Ramos PS, Kelly JA, Gray-McGuire C, Bruner GR, Leiran AN, Meyer CM, Namjou B, Espe KJ, Ortmann WA, Reichlin M, Langefeld CD, James JA, Gaffney PM, Behrens TW, Harley JB, Moser KL: **Familial aggregation and linkage analysis of autoantibody traits in pedigrees multiplex for systemic lupus erythematosus.** *Genes Immun* 2006, **7**(5):417-432.
50. Scofield RH, Bruner GR, Kelly JA, Kilpatrick J, Bacino D, Nath SK, Harley JB: **Thrombocytopenia identifies a severe familial phenotype of systemic lupus erythematosus and reveals genetic linkages at 1q22 and 11p13.** *Blood* 2003, **101**(3):992-997.
51. Tsao BP, Grossman JM, Riemekasten G, Strong N, Kalsi J, Wallace DJ, Chen CJ, Lau CS, Ginzler EM, Goldstein R, Kalunian KC, Harley JB, Arnett FC, Hahn BH, Cantor RM: **Familiality and co-occurrence of clinical features of systemic lupus erythematosus.** *Arthritis Rheum* 2002, **46**(10):2678-2685.

Table 1: SLE susceptibility loci included in the study with the representative SNPs, their Minor Allele Frequencies (MAF) in patients and controls, O.R. and P values^a

Locus	SNP	MAF % (n/N) ^b		O.R. (95% C.I.)	P-value ^c
		SLE patients	Controls		
<i>ITGAM</i>	<i>rs1143679</i>	23.0 (718/3116)	15.1 (521/3448)	1.71 (1.5-2.0)	8.1 x 10 ⁻¹⁵
<i>STAT4</i>	<i>rs7574865</i>	32.6 (1015/3118)	23.4 (808/3456)	1.53 (1.4-1.7)	6.0 x 10 ⁻¹³
<i>C8orf13-BLK</i>	<i>rs13277113</i>	30.9 (963/3116)	25.2 (868/3444)	1.32 (1.2-1.5)	2.0 x 10 ⁻⁶
<i>IRAK1-MECP2</i>	<i>rs17435</i>	27.0 (841/3120)	21.5 (743/3456)	1.27 (1.1-1.4)	2.5 x 10 ⁻⁵
<i>1q25.1</i>	<i>rs10798269</i>	27.1 (846/3122)	31.8 (1098/3452)	0.81 (0.7-0.9)	0.00019
<i>PXK</i>	<i>rs6445975</i>	26.7 (834/3118)	23.5 (813/3454)	1.24 (1.1-1.4)	0.00037
<i>TYK2</i>	<i>rs2304256</i>	23.3 (727/3116)	27.8 (960/3450)	0.82 (0.7-0.9)	0.00076
<i>BANK1</i>	<i>rs17266594</i>	25.0 (779/3122)	29.0 (1003/3456)	0.83 (0.7-0.9)	0.0017
<i>KIAA1542</i>	<i>rs4963128</i>	30.2 (939/3114)	34.0 (1173/3448)	0.84 (0.8-0.9)	0.0011

^a These analyses were reported in detail in Suarez-Gestal et al (22)

^b n = number of minor alleles, N = total number of alleles

^c Comparisons were done with logistic regression following a genetic additive model and including gender as covariate except for *IRAK1-MECP2*, which is in the X chromosome

Table 2: Clinical characteristics of the patients with SLE included in the study and the number of subjects with available information for each of them

Characteristic	Value	Number of subjects
Women (%)	90.2	1544
Age of disease onset (median, IQR ^a)	29 (21-40)	1340
Age of disease onset < 30 (%)	52.0	1340
Time of follow-up (median, IQR)	10 (6-17)	1159 ^b
Malar rash (%)	56.6	1296 ^b
Discoid rash (%)	18.1	1379
Photosensitivity (%)	53.6	1388
Oral ulcers (%)	28.1	1390
Arthritis (%)	80.6	1398
Serositis (%)	35.8	1392
Renal disorder (%)	41.0	1387
Neurologic disorder (%)	14.5	1305
Hematologic disorder (%)	74.2	1348
Immunologic disorder (%)	77.6	1383
Antinuclear Antibody (%)	98.5	1129 ^b

^a IQR = Inter-quartile range

^b Data from specific recruitment centres were excluded because the feature was underreported: the Czech Republic and Madrid for time of follow-up; Hungary for malar rash; and the Evangelismos Hospital in Greece and the Milan collections for ANA

Table 3: Comparison of the allele frequencies at nine SNP from SLE susceptibility loci between patients positive and negative for 11 clinical features. Only significant differences and results corresponding to previously reported associations are shown.

Locus	Feature	MAF % (n/N) ^a		O.R. (95% C.I.)	P	P _{corr}
		Positive patients	Negative patients			
<i>C8orf13-BLK</i>	Renal disorder	34.8 (395/1136)	28.0 (457/1632)	1.37 (1.2-1.6)	0.00017	0.0019
<i>IRAK1-MECP2</i>	Renal disorder	29.9 (320/1072)	25.2 (392/1556)	1.25 (1.1-1.5)	0.0063	0.06
<i>STAT4</i>	Age of onset <30	34.9 (486/1394)	31.1 (399/1282)	1.18 (1.0-1.4)	0.04	n.s. ^b
	Oral ulcers	30.1 (235/780)	33.7 (672/1996)	0.85 (0.7-1.0)	0.07	n.s.
	Renal disorder	32.9 (374/1136)	32.7 (535/1634)	1.01 (0.9-1.2)	n.s.	n.s.
	Immunologic d.	33.3 (715/2144)	30.7 (190/618)	1.13 (0.9-1.4)	n.s.	n.s.
<i>BANK1</i>	Oral ulcers	21.5 (168/780)	26.1 (521/2000)	0.78 (0.6-1.0)	0.01	n.s.
	Renal disorder	22.7 (258/1136)	26.2 (429/1638)	0.82 (0.7-1.0)	0.03	n.s.
<i>PXK</i>	Oral ulcers	23.7 (184/776)	27.9 (557/2000)	0.82 (0.7-1.0)	0.04	n.s.
<i>ITGAM</i>	Age of onset <30	24.7 (344/1392)	21.2 (272/1282)	1.22 (1.0-1.5)	0.03	n.s.
	Discoid rash	23.9 (119/498)	22.6 (509/2254)	1.08 (0.9-1.4)	n.s.	n.s.
	Renal disorder	23.7 (269/1134)	22.2 (363/1634)	1.09 (0.9-1.3)	n.s.	n.s.
	Immunologic d.	23.3 (498/2140)	21.5 (133/620)	1.11 (0.9-1.4)	n.s.	n.s.

^a MAF = Minor Allele Frequency, n = number of minor alleles, N = total number of alleles

^b n.s. = not significant

Table 4: Stratification of the significant associations from Table 3 by gender

Locus	Feature	Women		Men	
		O.R. (95% C.I.)	P	O.R. (95% C.I.)	P
<i>C8orf13-BLK</i>	Renal disorder	1.32 (1.1-1.6)	0.0014	1.84 (1.1-3.1)	0.023
<i>IRAK1-MECP2</i>	Renal disorder	1.22 (1.0-1.5)	0.029	2.15 (1.0-4.6)	0.05
<i>STAT4</i>	Age of onset <30	1.15 (1.0-1.3)	0.075	1.62 (0.9-2.9)	0.097
<i>BANK1</i>	Oral ulcers	0.78 (0.6-1.0)	0.018	0.79 (0.4-1.6)	0.5
	Renal disorder	0.79 (0.7-1.0)	0.018	1.03 (0.6-1.8)	0.9
<i>PXK</i>	Oral ulcers	0.79 (0.6-1.0)	0.024	1.13 (0.6-2.2)	0.71
<i>ITGAM</i>	Age of onset <30	1.22 (1.0-1.5)	0.04	1.19 (0.7-2.1)	0.55

Table 5. Combined analysis of data from current and previous studies. Only associations significant in any of the studies and allowing for case-only analysis were considered.

Locus	Feature	OR this study	OR previous study ^a	O.R.-comb. ^b (95% C.I.)	<i>P</i> _{comb.} ^b	<i>P</i> _{heter.} ^c	N
<i>C8orf13-BLK</i>	Renal disorder	1.37 (1.2-1.6)	0.73 (0.5-1.1)	-	-	0.0018	2165
<i>STAT4</i>	Age of onset <30	1.18 (1.0-1.4)	1.22 (1.0-1.4)	1.20 (1.1-1.3)	0.0017	0.78	2696
	Oral ulcers	0.85 (0.7-1.0)	0.80 (0.7-0.9)	0.82 (0.7-0.9)	0.0009	0.6	2783
	Immunologic d.	1.13 (0.9-1.4)	1.24 (1.0-1.5)	1.19 (1.0-1.4)	0.011	0.51	2777
	Renal disorder	1.01 (0.9-1.2)	1.23 (1.0-1.5)	1.07 (1.0-1.2)	0.14	0.061	2780
<i>ITGAM</i>	Renal disorder	1.09 (0.9-1.3)	1.39 (1.2-1.7)	1.23 (1.1-1.4)	0.0013 ^d	0.058	3098
	Discoid rash	1.08 (0.9-1.4)	1.27 (1.0-1.6)	1.16 (1.0-1.3)	0.029 ^d	0.24	2858
	Immunologic d.	1.11 (0.9-1.4)	1.30 (1.0-1.7)	1.18 (1.0-1.4)	0.04	0.35	3068

^a Previous studies were [15] for *C8orf13-BLK*, [9] for *STAT4* and [11] for *ITGAM*

^b Fixed-effect metaanalysis odds ratio (95 % confidence interval) and *P* value

^c *P*_{heter} = heterogeneity *P* value; N = total number of patients in the combined analysis

^d Random-effect metaanalysis gave non-significant results

Additional files provided with this submission:

Additional file 1: competinginterests 1392903058347436.doc, 44K
<http://arthritis-research.com/imedia/7827810763522699/supp1.doc>

Differences of Systemic Lupus Erythematosus Clinical Features between two Large European Population Subgroups

Marian Suarez-Gestal¹, Manuel Calaza¹, Torsten Witte², Chryssa Papasteriades³, Maurizio Marchini⁴, Sergio Migliaresi⁵, Attila Kovacs⁶, Josep Ordi Ros⁷, Marc Bijl⁸, Maria Jose Santos⁹, Sarka Ruzickova¹⁰, Rudolf Pullmann¹¹, Patricia Carreira¹², Fotini N. Skopouli¹³, Sandra D'Alfonso¹⁴, Gian Domenico Sebastiani¹⁵, Ana Suarez¹⁶, Francisco J Blanco¹⁷, Juan J Gomez-Reino^{1,18}, Antonio Gonzalez¹, and the **European Consortium of SLE DNA Collections**

ACADEMIC DEGREES:

M Suarez-Gestal, MSc; M Calaza, MSc; T Witte, MD; C Papasteriades, MD; M Marchini, MD; S Migliaresi, MD; A Kovacs, MD; J Ordi Ros, MD; M Bijl, MD; MJ Santos, MD; S Ruzickova, MD; R Pullmann, MD; P Carreira, MD; FN Skopouli, MD; S D'Alfonso, MD; GD Sebastiani, MD; A Suarez, MD; FJ Blanco, MD; JJ Gomez-Reino, MD; A Gonzalez, MD.

AFFILIATIONS:

¹ Laboratorio de Investigacion 10 and Rheumatology Unit, Instituto de Investigacion Sanitaria - Hospital Clinico Universitario de Santiago, 15706-Santiago de Compostela, Spain

² Division of Clinical Immunology, Department of Internal Medicine of the Hannover Medical School, D-30625 Hannover, Germany

³ Department of Histocompatibility and Immunology, Evangelismos Hospital, 10675-Athens, Greece

- ⁴ Clinical Immunology, University of Milan and Fondazione IRCCS Ospedale Maggiore Policlinico, Mangiagalli e Regina Elena, 20122- Milan, Italy
- ⁵ Rheumatology Unit, Second University of Naples, 81100-Naples, Italy
- ⁶ Department of Rheumatology, Hospital of Hungarian Railways, H-5000 Szolnok, Hungary
- ⁷ Internal Medicine, Research Laboratory in Autoimmune Diseases. Hospital Vall d'Hebron. 08035-Barcelona. Spain
- ⁸ Department of Rheumatology and Clinical Immunology, University Medical Center Groningen, 9713-Groningen, The Netherlands
- ⁹ Rheumatology Department, Hospital Garcia de Orta and Rheumatology Reseach Unit, Instituto de Medicina Molecular. Lisboa, Portugal
- ¹⁰ Institute of Biotechnology, Academy of Sciences of the Czech Republic, 128 50 Prague 2, Czech Republic
- ¹¹ Institute of Clinical Biochemistry, Martin Faculty Hospital, Jessenius Medical Faculty, Kollárova 2, 036 59 Martin, Slovakia
- ¹² Rheumatology Department. Hospital 12 de Octubre. 28041-Madrid. Spain; carreira@h12o.es
- ¹³ Pathophysiology Department, Athens University Medical School, Athens – 115 27, Greece
- ¹⁴ Dept. Medical Sciences and IRCAD , Eastern Piedmont University, 28100-Novara, Italy
- ¹⁵ Ospedale S. Camillo - Forlanini, U.O. Complessa di Reumatologia, 00151-Roma, Italy
- ¹⁶ Department of Functional Biology, Hospital Universitario Central de Asturias, Universidad de Oviedo, Oviedo 33006, Spain
- ¹⁷ Laboratorio de Investigación Osteoarticular y del Envejecimiento. Servicio de Reumatología. CH Universitario A Coruña. 15006-A Coruña. Spain
- ¹⁸ Department of Medicine, University of Santiago de Compostela, Spain

Corresponding author:

Antonio Gonzalez

Laboratorio de Investigacion 10

Hospital Clinico Universitario de Santiago

Travesia de Choupana sn.

15706-Santiago de Compostela

Spain

Tfn: 34 981 950 903

Fax: 34 981 951 068

Antonio.Gonzalez.Martinez.Pedrayo@sergas.es

anlugon@hotmail.com

Running title: European subgroups in SLE

FUNDING

MS-G is the recipient of a FPU pre-doctoral bursary of the Spanish Ministry of Education. This work was supported by Fondo de Investigacion Sanitaria of the Instituto de Salud Carlos III (Spain), grants [06/0620, 08/0744 and by RETICS Program, RD08/0075 (RIER)] that are partially financed by the Fondo Europeo de Desarrollo Regional program of the European Union, and by grants from the Xunta de Galicia and by BMBF KN Rheuma grant [C2.12 to TW].

CONFLICT OF INTERESTS

No conflict of interest has been declared by the authors.

ABBREVIATIONS

SLE: Systemic Lupus Erythematosus

ACR: American College of Rheumatology

CV: Coefficient of variation

ABSTRACT

We aimed to determine if Systemic Lupus Erythematosus (SLE) phenotypes differ between the two subgroups of Europeans identified in large genetic studies. Prevalence of ACR SLE classification criteria and mean age of disease onset were compared between 475 patients with SLE from the Northern subgroup and 1080 patients from the Southern subgroup. Arthritis was more prevalent in the Southern subgroup ($P = 1.7 \times 10^{-6}$). Photosensitivity was more prevalent in the Northern group ($P = 0.0006$). These differences were concordant with a previous study that analyzed European ancestry percentages in American patients with SLE, confirming a new level of phenotype heterogeneity of SLE inside one of the continental ethnic groups and suggesting the need of new levels of stratification in genetic and population studies.

INTRODUCTION

SLE clinical heterogeneity poses many challenges to clinical diagnosis, treatment and research¹¹. Genetic, environmental and socioeconomic factors seem to have a role, but our understanding of their effects is still very incomplete. Variation on severity and clinical phenotype between ethnic groups has been shown in multiple studies^{1, 4, 6, 8, 11-12}. Part of this variation is attributed to differences in the genetic component. A very interesting addition that strengthens the link between genetic factors and SLE clinical heterogeneity and extends the distinction to differences within a single ethnic group has been a recent report by Chung et al.³. In this study, distinction between Northern and Southern European ancestries, which are the 2 major European subgroups according to genome wide studies⁷, correlated with the prevalence of specific SLE features. Replication of these findings is necessary given the high variability of SLE clinical data between patient collections that makes this type of studies prone to irreproducibility^{1-2, 5-6, 10, 12}. It is also necessary given the important consequences for SLE research in Europeans that will need to consider separately the 2 subpopulations. Here, we have replicated 2 of the associations described by Chung et al: arthritis was more frequent in patients from the South of Europe and photosensitivity was more frequent in patients from the North.

METHODS

Patient data: Data from 1555 European SLE patients recruited at 16 centres from 9 different countries were analyzed⁹. They included the American College of Rheumatology (ACR) classification criteria, age of onset, time of follow-up and gender. All patients and controls gave their written informed consent and the study was approved by the relevant ethic committees.

Statistical analysis: Patients were divided in Northern and Southern subgroups according with their place of origin. We excluded from analysis the variables showing large variability within the 2 subgroups (Coefficient of variation (CV) ≥ 30 % between the centres in the Northern or Southern subpopulations). Comparison of the 2 subpopulations was done with logistic regression

incorporating patient gender and time of follow-up as covariates.

Founding sources: This work was supported by Fondo de Investigacion Sanitaria of the Instituto de Salud Carlos III, the Fondo Europeo de Desarrollo Regional program of the European Union, grants from the Xunta de Galicia and by BMBF KN Rheuma grant. This founding bodies had no contribution in designing the study, collecting, analyzing and interpreting the data and approving the final manuscript.

RESULTS

The Northern European subgroup was made of 475 patients with SLE from Germany, The Netherlands, Hungary, the Czech Republic and Slovakia. The Southern Europeans subgroup was made of 1080 patients from Spain, Italy, Greece and Portugal.

The first step in the analysis was to determine the variability of SLE clinical data within each of the 2 subgroups. This analysis was done separately for the centres corresponding to Southern and Northern European subpopulations (Table 1). Discoid rash, oral ulcers, renal disorder and neurologic disorder showed CV that were larger than 30 % in the Southern and Northern subgroups and were excluded from further analysis. Serositis, immunologic disorder, malar rash and photosensitivity showed CV of 20-25 %, whereas arthritis, hematologic disorder and antinuclear antibodies CV were below 20 % (Table 1). Variability in patient gender and age of disease onset was low. Time of follow-up was more variable in the Southern subgroup and significantly different between the 2 subgroups (Table 2). However, this difference in time of follow-up had not appreciable effect in the results except for serositis, which was less prevalent in the patients with shorter follow-up (not shown).

Only arthritis, photosensitivity and serositis were different between the patients from the Southern and Northern European subpopulations (Table 2). The differences in prevalence of arthritis and photosensitivity remained clearly significant after accounting for the number of test

performed ($P_{\text{corr}} = 0.000017$ and 0.0056 , respectively), but not the difference in serositis ($P_{\text{corr}} = 0.054$). In addition, the association with serositis disappeared when the difference in follow-up between the Northern and Southern subpopulations was taken into account ($P = 0.6$).

Notably, the differences in arthritis and photosensitivity were in the same direction previously reported in the European Americans³, with a prevalence in excess of 10.7 % in arthritis among the patients from the South and of 9.4 % in photosensitivity among the subjects from the Northern group.

DISCUSSION

Differences in the SLE phenotype between continental ethnic populations are widely acknowledged^{1, 4, 6, 8, 11-12}. However, there is only a report indicating that there are also differences between patients of European origin depending on their Northern or Southern ancestry³. Here, we have addressed this question and found 2 clear differences, in arthritis and photosensitivity, that are concordant with the study of Chung et al.³. This concordance is very striking because design of the 2 studies has been very different. The Chung et al. study included almost exclusively European Americans (98.0 %) whose ancestry was ascertained by analysis of ancestry informative polymorphisms. In contrast, we have used data from European patients living in Europe and their ancestry has been ascertained by their place of origin. This difference is important because each collection of patients is affected by environmental exposures and other local factors that could confound analysis, but these factors will be very different in the 2 studies. Therefore, we find very reassuring that the arthritis and photosensitivity results of the 2 studies were concordant.

The differences we have found could be indicative of underlying mechanisms whose elucidation will help understand SLE clinical heterogeneity. For example, photosensitivity excess in the Northern subpopulation could be related with a lighter skin colour^{3, 8}, which is genetically determined.

Other associations of the Chung et al. study were not replicated in our analysis³. We did not

found any difference between the 2 subgroups of Europeans in the prevalence of immunologic disorder and we did not compare discoid rash or renal disorder because these 2 features showed large variability within each of the subgroups in our data. High variability in the prevalence clinical features of SLE is very common between different collections of patients^{1-2, 5-6, 10, 12}. When it is present it is best to avoid comparisons because results are unreliable.

In summary, our study replicates the association of the Northern European subpopulation with more photosensitivity and less arthritis than the observed in the Southern subpopulation. Concordance between our study and Chung et al.³ identify SLE clinical features under genetic influence, advance in our understanding of SLE clinical heterogeneity, and call our attention to the need to account for European subpopulations in SLE studies.

ACKNOWLEDGMENTS:

Other contributors to the **European Consortium of SLE DNA Collections**: Myriam Liz, Laboratorio de Investigacion 10 and Rheumatology Unit, Instituto de Investigacion Sanitaria – Hospital Clinico Universitario de Santiago, 15706-Santiago de Compostela, Spain; Reinhold E Schmidt, Division of Clinical Immunology, Department of Internal Medicine of the Hannover Medical School, D-30625 Hannover, Germany; Iris Kappou-Rigatou, Department of Histocompatibility and Immunology, Evangelismos Hospital, 10676 Athens, Greece; Raffaella Scorza, Clinical Immunology, University of Milan and Fondazione IRCCS Ospedale Maggiore Policlinico, Mangiagalli e Regina Elena, 20122 Milan, Italy; Emöke Endereffy, Department of Paediatrics, Albert Szent-Györgyi Clinical Centre, University of Szeged, 6720 Hungary; Eva Balada, Internal Medicine, Research Laboratory in Autoimmune Diseases Hospital Vall d'Hebron, 08035 Barcelona, Spain; Cees G Kallenberg, Department of Rheumatology and Clinical Immunology, University Medical Center Groningen, 9713 Groningen, The Netherlands; Filipe Vinagre, Rheumatology Department, Hospital Garcia de Orta, Almada, (Portugal) and Rheumatology Research Unit, Instituto Medicina Molecular, Faculdade de Medicina da

Universidade de Lisboa, Portugal; Ctibor Dostal, Molecular Biology and Immunogenetics Dept, Institute of Rheumatology, 128 50 Prague 2, Czech Republic; Rudolf Pullmann Jr, Gerontology Research Center, National Institute on Aging, 5600 Nathan Shock Drive, Baltimore, Maryland 21224, USA; Maria Mavromati, Pathophysiology Department, Athens University Medical School, Athens 115 27, Greece; Nadia Barizzzone, Dept Medical Sciences and IRCAD , Eastern Piedmont University, 28100 Novara, Italy; Carmen Gutierrez, Department of Functional Biology, Hospital Universitario Central de Asturias, Universidad de Oviedo, Oviedo 33006, Spain; Ignacio Rego, Laboratorio de Investigación Osteoarticular y del Envejecimiento. Servicio de Reumatología. CH Universitario A Coruña. 15006-A Coruña. Spain.

REFERENCES

1. Alarcon GS, McGwin G, Jr., Petri M, Reveille JD, Ramsey-Goldman R, Kimberly RP. Baseline characteristics of a multiethnic lupus cohort: PROFILE. *Lupus* 2002;11: 95-101.
2. Cervera R, Abarca-Costalago M, Abramovicz D, Allegri F, Annunziata P, Aydintug AO, Bacarelli MR, Bellisai F, Bernardino I, Biernat-Kaluza E, Blockmans D, Boki K, Bracci L, Campanella V, Camps MT, Carcassi C, Cattaneo R, Cauli A, Chwalinska-Sadowska H, Contu L, Cosyns JP, Danieli MG, D DC, Depresseux G, Direskeneli H, Domenech I, Espinosa G, Fernandez-Nebro A, Ferrara GB, Font J, Frutos MA, Galeazzi M, Garcia-Carrasco M, Garcia Iglesias MF, Garcia-Tobaruela A, George J, Gil A, Gonzalez-Santos P, Grana M, Gul A, Haga HJ, de Haro-Liger M, Houssiau F, Hughes GR, Ingelmo M, Jedryka-Goral A, Khamashta MA, Lavilla P, Levi Y, Lopez-Dulpa M, Lopez-Soto A, Malykova H, Marcolongo R, Mathieu A, Morozzi G, Nicolopoulou N, Papasteriades C, Passiu G, Perello I, Petera P, Petrovic R, Piette JC, Pintado V, de Pita O, Popovic R, Pucci G, Puudu P, de Ramon E, Ramos-Casals M, Rodriguez-Andreu J, Ruiz-Irastorza G, Sanchez-Lora J, Sanna G, Scorza R, Sebastiani GD, Sherer Y, Shoenfeld Y, Simpatico A, Sinico RA, Smolen J, Tincani A, Tokgoz G, Urbano-Marquez A, Vasconcelos C, Vazquez JJ, Veronesi J, Vianna J, Vivancos J. Systemic lupus erythematosus in Europe at the change of the millennium: lessons from the "Euro-Lupus Project". *Autoimmun Rev* 2006;5: 180-6.
3. Chung SA, Tian C, Taylor KE, Lee AT, Ortmann WA, Hom G, Graham RR, Nititham J, Kelly JA, Morrissey J, Wu H, Yin H, Alarcon-Riquelme ME, Tsao BP, Harley JB, Gaffney PM, Moser KL, Manzi S, Petri M, Gregersen PK, Langefeld CD, Behrens TW, Seldin MF, Criswell LA. European population substructure is associated with mucocutaneous manifestations and autoantibody production in systemic lupus erythematosus. *Arthritis Rheum* 2009;60: 2448-56.
4. Fernandez M, Alarcon GS, Calvo-Alen J, Andrade R, McGwin G, Jr., Vila LM, Reveille JD. A multiethnic, multicenter cohort of patients with systemic lupus erythematosus (SLE) as a model for the study of ethnic disparities in SLE. *Arthritis Rheum* 2007;57: 576-84.
5. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, Lee AT, Chung SA,

- Ferreira RC, Pant PV, Ballinger DG, Kosoy R, Demirci FY, Kamboh MI, Kao AH, Tian C, Gunnarsson I, Bengtsson AA, Rantapaa-Dahlqvist S, Petri M, Manzi S, Seldin MF, Ronnblom L, Syvanen AC, Criswell LA, Gregersen PK, Behrens TW. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. *N Engl J Med* 2008;358: 900-9.
6. Petri M. The effect of race on the presentation and course of SLE in the United States. *Arthritis and Rheumatism* 1997;40: 785-785.
7. Seldin MF, Shigeta R, Villoslada P, Selmi C, Tuomilehto J, Silva G, Belmont JW, Klareskog L, Gregersen PK. European population substructure: clustering of northern and southern populations. *PLoS Genet* 2006;2: e143.
8. Sestak AL, Nath SK, Kelly JA, Bruner GR, James JA, Harley JB. Patients with familial and sporadic onset SLE have similar clinical profiles but vary profoundly by race. *Lupus* 2008;17: 1004-9.
9. Suarez-Gestal M, Calaza M, Endreffy E, Pullmann R, Ordi-Ros J, Domenico Sebastiani G, Ruzickova S, Jose Santos M, Papasteriades C, Marchini M, Skopouli FN, Suarez A, Blanco FJ, D'Alfonso S, Bijl M, Carreira P, Witte T, Migliaresi S, Gomez-Reino JJ, Gonzalez A. Replication of recently identified systemic lupus erythematosus genetic associations: a case-control study. *Arthritis Res Ther* 2009;11: R69.
10. Taylor KE, Remmers EF, Lee AT, Ortmann WA, Plenge RM, Tian C, Chung SA, Nititham J, Hom G, Kao AH, Demirci FY, Kamboh MI, Petri M, Manzi S, Kastner DL, Seldin MF, Gregersen PK, Behrens TW, Criswell LA. Specificity of the STAT4 genetic association for severe disease manifestations of systemic lupus erythematosus. *PLoS Genet* 2008;4: e1000084.
11. Wallace DJ, Hahn B, Dubois EL. Dubois' lupus erythematosus. 7th edn. Philadelphia: Lippincott Williams & Wilkin, 2007.
12. Wang F, Wang CL, Tan CT, Manivasagar M. Systemic lupus erythematosus in Malaysia: a study of 539 patients and comparison of prevalence and disease expression in different racial and gender groups. *Lupus* 1997;6: 248-53.

Table 1: Clinical features of SLE patients and variability of their values within the European Northern and Southern subgroups.

	Value	N *	CV (%)	
			Northern †	Southern †
Women (%)	90.3	1543	3.1	3.3
age of onset (mean, years)	31.1	1349	6.3	10.6
time of follow-up (mean, years)	11.9	1168	12.6	22.8
Malar rash (%) ‡	56.5	1305	24.8	19.8
Discoid rash (%)	18.1	1388	48.8	59.2
Photosensitivity (%)	53.4	1397	17.6	23.0
Oral ulcers (%)	28.2	1399	56.0	63.1
Arthritis (%)	80.7	1407	15.8	10.4
Serositis (%)	35.8	1401	9.9	26.2
Renal disorder (%)	41.0	1396	32.5	38.9
Neurologic disorder (%)	14.6	1314	66.4	38.4
Hematologic disorder (%)	74.3	1356	15.0	18.8
Immunologic disorder (%)	77.7	1392	22.1	23.0
Antinuclear antibody (%) ‡	98.7	1138	1.8	2.1

* Number of patients used for each analysis

† Northern subgroup: Germany, the Netherlands, Hungary, the Czech Republic and Slovakia. Southern subgroup: Spain, Italy, Greece and Portugal.

‡ Some data from all patients of specific recruitment centres were excluded because the feature was underreported: Hungary for malar rash; and the Evangelismos Hospital in Greece and the Milan collections for antinuclear antibody

Table 2: Comparison of clinical features between the Northern and Southern subpopulations of European patients with SLE

	Value		P-value *	O.R. (95 % C.I.) †
	Northern	Southern		
Women (%)	88.6	91.0	0.14	
age of onset (mean, years)	31.9	30.8	0.18	
time of follow-up (mean, years)	14.2	10.7	5.3×10^{-14}	
Malar rash (%)	56.3	55.6	0.17	1.01 (0.9-1.2)
Photosensitivity (%)	59.7	50.3	0.00056	1.49 (1.2-1.9)
Arthritis (%)	73.6	84.3	0.0000017	0.51 (0.4-0.7)
Serositis (%)	39.5	33.8	0.0055 ‡	1.29 (1.1-1.7)
Hematologic disorder (%)	75.5	73.7	0.12	1.11 (0.9-1.4)
Immunologic disorder (%)	77.2	77.9	0.68	0.95 (0.7-1.2)
Antinuclear antibody (%)	98.6	98.5	0.92	1.06 (0.3-3.3)

* Comparison done with logistic regression including patient gender as covariate

† Odds ratios and their 95 % confidence intervals were calculated taking as reference the Southern subpopulation of patients

‡ This difference did not persist after considering differences in follow-up

4. Componente genético común a LES y AR

Rheumatoid Arthritis Does Not Share Most of the Newly Identified Systemic Lupus Erythematosus Genetic Factors

Marian Suarez-Gestal,¹ Manuel Calaza,¹ Rebeca Dieguez-Gonzalez,¹ Eva Perez-Pampin,¹ Jose Luis Pablos,² Federico Navarro,³ Javier Narvaez,⁴ Jose Luis Marengo,⁵ Gabriel Herrero-Beaumont,⁶ Benjamin Fernandez-Gutierrez,⁷ Jose Ramon Lamas,⁷ Arturo Rodriguez de la Serna,⁸ Ana Maria Ortiz,⁹ Luis Carreño,¹⁰ Juan D. Cañete,¹¹ Rafael Caliz,¹² Francisco J. Blanco,¹³ Alejandro Balsa,¹⁴ Juan J. Gomez-Reino,¹⁵ and Antonio Gonzalez¹

Objective. Rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) share some genetic factors such as *HLA*, *PTPN22*, *STAT4*, and *6q23*. The

aim of this study was to determine whether 9 other SLE genetic factors are also implicated in RA susceptibility.

Methods. A characteristic single-nucleotide polymorphism (SNP) in each of 9 genetic factors, *ITGAM* (rs1143679), *C8orf13-BLK* (rs13277113), *TYK2* (rs2304256), *Iq25.1* (rs10798269), *PXK* (rs6445975), *KIAA1542* (rs4963128), *MECP2* (rs17435), *BANK1* (rs17266594), and *LY9* (rs509749), was studied in 1,635 patients with RA and 1,906 control subjects from Spain. The rs7574865 SNP in *STAT4* was also included. Analyses were conducted globally and after stratification by sex and clinical features (anti-cyclic citrullinated peptide and rheumatoid factor, shared epitope, rheumatoid nodules, radiographic changes, sicca syndrome, and pneumonitis).

Results. No association was observed between RA and any of the 9 newly identified SLE genetic factors. A meta-analysis using previous data was consistent with these results. In addition, there were no significant differences between individuals with and those without each of the clinical features analyzed, except the frequency of the minor allele in the *C8orf13-BLK* locus that was decreased in patients with sicca syndrome (14.6% versus 22.4% in controls; $P = 0.003$).

Conclusion. None of the 9 recently identified SLE risk factors showed association with RA. Therefore, common genetic factors affecting the pathogenesis of these 2 disorders seem to be limited, revealing that the genetic component contributes to the different expression of these diseases.

Rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) share a complex etiology encom-

Supported by the Instituto de Salud Carlos III of the Spanish Health Ministry (grants PI06/0620 and PI08/0744) and the Xunta de Galicia (grant 07PXIB918091PR). The Instituto de Salud Carlos III Research Network provided support for many of the collaborating research teams (grant RD08/0075).

¹Marian Suarez-Gestal, MSc, Manuel Calaza, MSc, Rebeca Dieguez-Gonzalez, PhD, Eva Perez-Pampin, MD, Antonio Gonzalez, MD, PhD: Hospital Clínico Universitario de Santiago, Santiago de Compostela, Spain; ²Jose Luis Pablos, MD: Hospital 12 de Octubre, Madrid, Spain; ³Federico Navarro, MD: Hospital Universitario Virgen Macarena, Seville, Spain; ⁴Javier Narvaez, MD, PhD: Hospital Universitario de Bellvitge, Barcelona, Spain; ⁵Jose Luis Marengo, MD: Hospital Universitario de Valme, Seville, Spain; ⁶Gabriel Herrero-Beaumont, MD: Fundacion Jimenez Diaz, Madrid, Spain; ⁷Benjamin Fernandez-Gutierrez, MD, PhD, José Ramón Lamas, PhD: Hospital Clínico San Carlos, Madrid, Spain; ⁸Arturo Rodriguez de la Serna, MD: Hospital Santa Creu i Sant Pau, Barcelona, Spain; ⁹Ana Maria Ortiz, MD, PhD: Hospital Universitario de la Princesa, Madrid, Spain; ¹⁰Luis Carreño, MD: Hospital Universitario Gregorio Marañón, Madrid, Spain; ¹¹Juan D. Cañete, MD, PhD: Hospital Clinicico and Institut d'Investigacions Biomèdiques August Pi I Sunyer, Barcelona, Spain; ¹²Rafael Caliz, MD: Hospital Universitario Virgen de las Nieves, Granada, Spain; ¹³Francisco J. Blanco, MD: Hospital Universitario Juan Canalejo, Coruña, Spain; ¹⁴Alejandro Balsa, MD: Hospital La Paz, Madrid, Spain; ¹⁵Juan J. Gomez-Reino, MD, PhD: Hospital Clínico Universitario de Santiago and Universitario de Santiago de Compostela, Santiago de Compostela, Spain.

Dr. Fernandez-Gutierrez has received speaking fees from Bristol-Myers Squibb (less than \$10,000). Dr. Balsa has received consulting fees, speaking fees, and/or honoraria from Abbott (less than \$10,000).

Address correspondence and reprint requests to Antonio Gonzalez, MD, PhD, Laboratorio de Investigacion 10, Hospital Clínico Universitario de Santiago, Travesía de Choupana s/n, 15706 Santiago de Compostela, Spain. E-mail: Antonio.Gonzalez.Martinez.Pedrayo@sergas.es.

Submitted for publication March 9, 2009; accepted in revised form May 17, 2009.

passing genetic, environmental, and stochastic components (1,2). Loss of tolerance to self antigens, which leads to stimulation of lymphocytes and other immune cells, release of cytokines, activation of complement, and the production of autoantibodies, contributes to the pathogenesis of both RA and SLE (3,4). RA and SLE also share genetic factors such as those in the *HLA*, *PTPN22*, *STAT4*, and *6q23* loci (5–8); however, their respective clinical phenotypes are clearly different. RA is characterized by symmetric erosive arthritis of the peripheral joints with pannus formation that is chronic and progressive, which causes a characteristic pattern of pain, rigidity, and deformities. In contrast, arthritis is only one among the diverse array of clinical manifestations of SLE and typically is not erosive, does not have a progressive course, and most often is not symmetric (9). Malar rash, photosensitivity, serositis, nephritis, peripheral and central nerve system disease, and cytopenia are indicative of SLE but not RA. The 2 clinical syndromes overlap only in some exceptional patients (a clinical situation that is referred to as *rheupus* by many investigators). Therefore, the 2 diseases should have differential etiologic factors. Such etiologic factors could be genetic, with some being specific for SLE and others being specific for RA. Alternatively, the differential factors could be environmental or stochastic.

A common genetic component of RA, SLE, and other autoimmune diseases has been hypothesized (10,11). This hypothesis was triggered by the observation of an increased concurrence of several autoimmune diseases in members of the same family (12,13) and was strongly reinforced by the discovery of the role of HLA alleles in most autoimmune diseases (14) and, later, by genetics studies in families showing that linkage to different autoimmune diseases clustered in overlapping loci (15). Similar findings were also observed in animal models of autoimmune diseases (16). More recently, some genetic factors with a shared effect in many of these diseases have been identified. The most generally shared is *PTPN22*, but other examples include *STAT4*, the *6q23* locus, and possibly *CTLA4* and *IRF5* (5–8,17,18). Now, thanks to genome-wide association studies, we are beginning to have more comprehensive knowledge of the genetic component of these diseases. The identification of more genetic factors will allow us to determine whether those that are shared by several diseases are the rule or the exception. As a consequence, we will make advances in the identification of differential disease factors.

At the time when this study was planned, the number of recently identified genetic factors in SLE was

Table 1. Clinical characteristics of the 1,635 patients with rheumatoid arthritis*

Female sex	75.0
Age at disease onset, median (interquartile range) years	48 (37–57)
Morning stiffness	96.1
Arthritis in ≥ 3 joint areas	99.7
Arthritis in hand joints	99.0
Symmetric arthritis	99.3
Rheumatoid nodules (n = 1,291)	20.1
Rheumatoid factor (n = 1,504)	72.7
Erosions (n = 1,502)	70.4
Sicca syndrome (n = 903)	14.8
Interstitial pneumonitis (n = 1,520)	2.8
Shared epitope carrier (n = 594)	54.7
Anti-citrullinated protein antibody positive (n = 655)	66.1

* Except where indicated otherwise, values are the percent. When data were not available for all patients, the number of patients for whom data were available is shown.

larger than the number of such factors in RA. Therefore, we examined the role in RA of a representative SNP for each of 10 recently identified SLE genetic factors: *STAT4* (which is already known to be shared by SLE and RA), *ITGAM*, *C8orf13-BLK*, *PXK*, *TYK2*, *KIAA1542*, *1q25.1*, *BANK1*, *MECP2*, and *LY9* (6). All except *LY9* have consistently been associated with SLE (19). However, only the SNP in *STAT4* was associated with RA in our current comparison of more than 1,600 patients and 1,900 healthy control subjects. A meta-analysis including previous data led to similar results. Therefore, it seems that RA does not share most of the new SLE genetic factors, indicating that the genetic component is a significant contributor to the different expression of these 2 diseases.

PATIENTS AND METHODS

Sample collection. We used DNA samples from 1,635 patients with RA and 1,906 healthy control subjects of Spanish ancestry (281 of the control subjects were previously described by our group in a study of SLE susceptibility [19]). All patients with RA met the 1987 revised American College of Rheumatology (formerly, the American Rheumatism Association) classification criteria (20). The clinical characteristics of the patients are shown in Table 1. Patients and control subjects gave written informed consent to participate in the study. Sample collection was approved by the respective ethics committees.

SNP genotyping. A strongly SLE-associated SNP for each of 10 recently identified SLE genetic factors was selected (Table 2). The 10 SNPs were amplified in a single polymerase chain reaction (PCR) (Qiagen Multiplex PCR Kit; Qiagen, Chatsworth, CA) with 20 ng of genomic DNA and 0.2 μ M of each primer. PCR products were purified by digestion with exonuclease I (Epicentre Technologies, Madison, WI) and shrimp alkaline phosphatase (SAP; GE Healthcare, Barcelona, Spain). Purified PCR products were included in a single-base

Table 2. Allele frequencies for the 10 SNPs identifying SLE genetic factors*

SNP (locus)	Minor allele frequency, no. (%)		
	RA patients	Control subjects	OR (95% CI)
rs7574865 (<i>STAT4</i>)	665/2,768 (24.0)	784/3,612 (21.7)	1.14 (1.0–1.3)†
rs1143679 (<i>ITGAM</i>)	561/3,264 (17.2)	608/3,808 (16.0)	1.09 (1.0–1.2)
rs13277113 (<i>C8orf13–BLK</i>)	757/3,268 (23.2)	852/3,810 (22.4)	1.05 (0.9–1.2)
rs2304256 (<i>TYK2</i>)	842/3,262 (25.8)	1,025/3,804 (26.9)	0.94 (0.8–1.0)
rs10798269 (<i>Iq25.1</i>)	1,042/3,268 (31.9)	1,285/3,812 (33.7)	0.92 (0.8–1.0)
rs4963128 (<i>KIAA1542</i>)	1,013/3,264 (31.0)	1,214/3,812 (31.8)	0.96 (0.9–1.1)
rs6445975 (<i>PXX</i>)	790/3,270 (24.2)	851/3,814 (22.3)	1.11 (1.0–1.2)
rs17266594 (<i>BANK1</i>)	773/2,768 (27.9)	987/3,612 (27.3)	1.03 (0.9–1.2)
rs509749 (<i>LY9</i>)	1,530/3,270 (46.8)	1,777/3,810 (46.6)	1.03 (0.9–1.2)
rs17435 (<i>MECP2</i>)‡			
Women	461/2,252 (20.5)	445/2,224 (20.0)	1.05 (0.9–1.2)
Men	71/372 (19.1)	100/622 (16.1)	1.23 (0.9–1.7)

* SLE = systemic lupus erythematosus; RA = rheumatoid arthritis; OR = odds ratio; 95% CI = 95% confidence interval.

† $P = 0.028$.

‡ This single-nucleotide polymorphism (SNP), which is in chromosome X, was assessed by analysis of allelic frequencies in women and by carrier analysis in men.

extension reaction with the SNaPshot Multiplex Kit (Applied Biosystems, Foster City, CA) and specific probes. After a second purification with SAP (GE Healthcare), samples were analyzed in the ABI Prism 3130xl Genetic Analyzer (Applied Biosystems), and genotypes were assigned using GeneMapper software (Applied Biosystems). (Primer and oligonucleotide sequences are available from the corresponding author.)

Statistical analysis. Tests for Hardy-Weinberg equilibrium in control samples were done with Haploview (21), using a threshold of 0.05, without correction for multiple testing. Other statistical analyses were done using a customized version of the Statistica 7.0 program (StatSoft, Tulsa, OK). Comparison of cases and controls was done with allele frequency data using chi-square tests on 2×2 contingency tables. These analyses were also conducted after stratifying the samples by sex and available clinical data. Results from the stratified analyses were subjected to Bonferroni correction. Univariate logistic regression models were used to test the fit to the data of additive, recessive, and dominant genetic models. Statistical power was estimated using “power and sample size calculations” software (22). Meta-analysis was done with our data and data from previous reports, using a fixed-effects model with R software (<http://www.r-project.org/>). Heterogeneity of the effect size between studies was assessed with the I^2 statistic for inconsistency and Cochran’s Q statistic. Samples from Hospital Universitario La Paz and Hospital Clinico San Carlos have already been used for the analysis of 2 genetic factors included in this project: *BANK1* and *STAT4*. To avoid data duplication, we excluded these samples from the relevant analysis.

RESULTS

A total of 1,635 patients with RA and 1,906 control subjects of Spanish ancestry were available for study. The genotyping call rate across the 10 SNPs was

99.9%, and genotypes in the control subjects were in Hardy-Weinberg equilibrium.

We observed association of the *STAT4* SNP rs7574865 with RA (odds ratio [OR] 1.14, 95% confidence interval CI 1.0–1.3, $P = 0.028$) but no significant differences between patients with RA and control subjects for the other 8 autosomal SNPs or for the *MECP2* SNP, which is located in the X chromosome and was analyzed separately in female and male individuals (Table 2). All genotype comparisons were concordant with an additive genetic model and gave results similar to the allele comparisons (results not shown). The power of these analyses was enough to detect genetic factors showing an OR between 1.20 for the *ITGAM* SNP and 1.15 for the *LY9* SNP ($\alpha = 0.05$ and $1 - \beta = 0.80$), which were the SNPs with minimum and maximum power, respectively.

Analysis was also done after stratification of samples according to sex and clinical data, which included anti-cyclic citrullinated peptide antibodies, rheumatoid factor, the shared epitope, rheumatoid nodules, radiographic changes, sicca symptoms, and interstitial pneumonitis. Most of these comparisons showed similar results across strata (detailed data are available from the corresponding author). The most remarkable difference was the decreased frequency of the minor allele of rs13277113 in the *C8orf13–BLK* locus among patients with RA and sicca symptoms (14.6% versus 24.2% in patients without these symptoms [$P = 0.0005$] and versus 22.4% in controls [$P = 0.003$]) (Table 3). These differ-

Table 3. Significant differences in stratified analyses of patients with RA, according to sex and clinical features*

SNP (locus)	Comparisons between patient strata				Comparisons with controls		
	MAF, no. (%)	OR (95% CI)	P	$P_{\text{corr}}^{\dagger}$	OR (95% CI)	P	$P_{\text{corr}}^{\dagger}$
rs6445975 (<i>PXK</i>)							
Women	541/2,252 (24.0)	1.10 (0.9–1.3)	NS	NS	1.19 (1.0–1.4)	0.015	NS
Men	168/752 (22.3)				0.87 (0.7–1.1)	NS	NS
rs13277113 (<i>C8orf13–BLK</i>)							
Sicca syndrome	39/268 (14.6)	0.53 (0.4–0.8)	0.0005	0.008	0.59 (0.4–0.8)	0.003	0.04
No sicca syndrome	356/1472 (24.2)				1.11 (1.0–1.3)	0.16	NS
rs509749 (<i>LY9</i>)							
Shared epitope	330/650 (50.8)	1.41 (1.1–1.8)	0.003	0.05	1.18 (1.0–1.4)	0.05	NS
No shared epitope	227/538 (42.2)				0.84 (0.7–1.0)	0.05	NS
rs10798269 (<i>Iq25.1</i>)							
RF positive	673/2,184 (30.8)	0.85 (0.7–1.0)	0.06	NS	0.88 (0.8–1.0)	0.02	NS
RF negative	283/822 (34.4)				1.03 (0.9–1.2)	NS	NS
rs2304256 (<i>TYK2</i>)							
Erosions	565/2,108 (26.8)	1.19 (1.0–1.4)	0.06	NS	0.99 (0.9–1.1)	NS	NS
No erosions	209/888 (23.5)				0.83 (0.7–1.0)	0.04	NS

* The presence of erosions was defined radiographically. RA = rheumatoid arthritis; SNP = single-nucleotide polymorphism; MAF = minor allele frequency; OR = odds ratio; 95% CI = 95% confidence interval; NS = not significant; RF = rheumatoid factor.

\dagger Bonferroni correction.

ences remained significant after correction for the number of tests performed (Table 3). However, this result should be taken with caution given the small number of patients in this group ($n = 134$) and association with a different allele than that in patients with SLE. There were other differences between strata (Table 3), but they were significant in the analyses only previous to Bonferroni correction.

Finally, we combined our results with previous data including data available only as supplementary information and data imputed from the available genotypes but not directly tested in previous studies. These analyses included, at least, an additional data set for all of the SNPs except *MECP2* SNP rs17435 and *ITGAM*

SNP rs1143679, for which no previous information was found. The meta-analysis confirmed the results observed in our sample collection: the one SNP that is clearly associated with RA is the *STAT4* SNP (Table 4), for which >35,000 subjects have already been studied. Results for this SNP showed significant heterogeneity of effects between studies ($I^2 = 53.0\%$, P for Cochran's Q statistic [P_Q] = 0.007) that was not attributable to the inclusion of our results ($I^2 = 53.2\%$, $P_Q = 0.008$). This heterogeneity prevents a sound estimate of the summary effect size, but it does not question association of the *STAT4* SNP with RA, because all studies showed an OR of >1.0 for this SNP. Two other SNPs showed borderline association with RA: the SNP in the *C8orf13–BLK*

Table 4. Combined analysis of allele frequency differences for SLE-associated SNPs obtained in the current and previous studies (including imputed genotypes) comparing patients with RA and controls*

SNP (locus)	No. of RA patients	No. of controls	OR (95% CI)	P	Ref.
rs7574865 (<i>STAT4</i>) \dagger	17,565	18,315	1.25 (1.2–1.3)	1.1×10^{-34}	31,38–46‡
rs7574865 (<i>STAT4</i>)§	16,181	16,509	1.26 (1.2–1.3)	2.9×10^{-34}	31,38–46‡
rs13277113 (<i>C8orf13–BLK</i>)	3,494	4,562	1.09 (1.0–1.2)	0.03	31
rs2304256 (<i>TYK2</i>)	3,161	2,763	0.94 (0.9–1.0)	NS	18
rs10798269 (<i>Iq25.1</i>)	5,006	6,694	0.97 (0.9–1.0)	NS	31,46
rs4963128 (<i>KLAAL542</i>)	5,014	6,694	1.01 (1.0–1.1)	NS	31,46
rs6445975 (<i>PXK</i>)	5,017	6,695	1.06 (1.0–1.1)	0.04	31,46
rs17266594 (<i>BANK1</i>)	3,550	3,960	0.97 (0.9–1.0)	NS	37
rs509749 (<i>LY9</i>)	5,014	6,691	1.04 (1.0–1.1)	NS	31,46

* SLE = systemic lupus erythematosus; SNPs = single-nucleotide polymorphisms; RA = rheumatoid arthritis; OR = odds ratio; 95% CI = 95% confidence interval; NS = not significant.

\dagger Analysis of heterogeneity between studies was possible only for *STAT4*, for which significant heterogeneity was observed ($I^2 = 53.0\%$, P for Cochran's Q statistic [P_Q] = 0.007).

‡ Data from refs. 39 and 46 overlap, and only the largest set from ref. 39 was used here.

§ $I^2 = 53.2\%$, $P_Q = 0.008$, not including data from the present study.

locus for which ~8,000 subjects have been analyzed ($P = 0.03$) and the SNP in *PXK*, with data from >11,000 subjects ($P = 0.04$).

DISCUSSION

During the last decades, evidence supporting a common genetic component for a diversity of autoimmune diseases has been continuously investigated (10,11,15,16,23). The genetic factors that were associated with a certain disease were considered good candidates for the other diseases in the group, and they were systematically examined. This practice could have introduced a bias supporting the commonality hypothesis, because extensive investigation of the same factors in multiple diseases could have led to some associations just by chance. This is likely in the context of the lack of reproducibility that until recently characterized genetics studies of complex diseases (24). Accordingly, conflicting results have been the rule. Nowadays, genetics studies have achieved a high degree of reproducibility due to a significant increase in sample size and to more strict thresholds for claiming association (25). In addition, association across the genome has been analyzed in an unbiased manner (25). These developments gave us the opportunity to address the question of genetic commonality between RA and SLE with a new set of solidly established genetic factors.

The results provide evidence against widespread sharing of genetic factors between RA and SLE. Of the 9 newly examined SLE genetic factors, none was found to be associated with RA. This lack of association persisted after a meta-analysis including previous data, which makes it unlikely that any of the 9 representative SNPs have an effect in RA similar to that observed in SLE.

In retrospect, it could be contended that our findings were expected, given the identification of these genetic factors in the SLE genome-wide association study (26–28) and not in the RA genome-wide association study (29–31). However, the lack of significant association in a genome-wide association study does not mean the lack of any association, because genome-wide association studies apply strict thresholds for significance, and coverage of the different SLE and RA genome-wide association studies was not identical (i.e., only 1 of the 10 SNPs studied here was included in the largest RA genome-wide association study [31]). In addition, 3 SLE genetic factors explored here, *TYK2* (32), *MECP2* (33), and *LY9* (34), were not found in the SLE genome-wide association study (although the first 2

factors have been independently confirmed [19,35]). Therefore, our specific analysis together with the meta-analysis including previous data allow for a more conclusive assessment that reinforces the SLE specificity of these genetic factors in relation to RA.

Our study has enough power to detect effect sizes corresponding to an OR between 1.15 and 1.20 for the SNPs in autosomes and 1.24 for the *MECP2* SNP in chromosome X. This is in the range of the weakest detectable effects by analysis of the largest RA sample collections, such as the North American Rheumatoid Arthritis Consortium family collection and the Epidemiological Investigation of Rheumatoid Arthritis (36). Combining our results with those from previous studies (including imputed data) increased the power. Two weak association signals with ORs of 1.09 and 1.06 were detected in the *C8orf13–BLK* and *PXK* loci, respectively. It is therefore possible that SNP effects below these sizes could have escaped our detection. In this case, the eventual RA-associated SNPs would be of lower significance for RA than for SLE disease, in which the reported ORs for these SNPs have been larger than 1.19, with the exception of single reports that did not replicate the *LY9* and *TYK2* associations (19,35). It is also possible that different polymorphisms in the same gene are involved in different diseases, as indicated by a recent report of association of *BANK1* SNPs with RA in a study that did not show association with the rs17266594 SNP that we examined here (37).

Another possibility we have addressed is that the sharing of genetic factors is restricted to specific subgroups of patients with RA, perhaps with more similarity with SLE. This is what happens with polymorphisms in *IRF5* that are associated with SLE and with a not-fully-defined subgroup of patients with RA (17,18). However, our analysis did not identify any clear association of this type. There were some association signals in the stratified analyses, but only differences in the *C8orf13–BLK* SNP in patients with sicca syndrome remained significant after correction. Unfortunately, no stratified data were available in previous reports that would allow us to draw more sound conclusions by increasing statistical power through meta-analysis.

In summary, the genetic factors in the *HLA*, *PTPN22*, *6q23*, *IRF5*, and *STAT4* loci are shared by SLE, RA, and other immune-mediated diseases (5,6). However, none of the other 9 recently identified SLE genetic factors examined here showed association with RA. These results indicate that the genetic component of each of these 2 diseases contributes significantly to the different disease phenotypes. This is an idea that is not

new but that lacked experimental support (10). The magnitude of the differential component cannot be estimated yet, because a large fraction of the heritability of RA and SLE has not been defined. More precise estimates will be possible once the causal variants of each of the genetic factors are identified. It is reasonable to expect that further genetic research will provide a clearer definition of what molecules and pathways are shared or specific in the pathogenesis of RA and SLE, allowing for better management of these 2 rheumatic diseases.

ACKNOWLEDGMENT

We thank the sample donors for their generous collaboration.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Gonzalez had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Suarez-Gestal, Gonzalez.

Acquisition of data. Dieguez-Gonzalez, Perez-Pampin, Pablos, Navarro, Narvaez, Marenco, Herrero-Beaumont, Fernandez-Gutierrez, Lamas, Rodriguez de la Serna, Ortiz, Carreño, Cañete, Caliz, Blanco, Balsa, Gomez-Reino, Gonzalez.

Analysis and interpretation of data. Suarez-Gestal, Calaza, Gomez-Reino, Gonzalez.

REFERENCES

1. Gregersen PK. Genetics of rheumatoid arthritis: confronting complexity. *Arthritis Res* 1999;1:37-44.
2. Nath SK, Kilpatrick J, Harley JB. Genetics of human systemic lupus erythematosus: the emerging picture. *Curr Opin Immunol* 2004;16:794-800.
3. Ronnblom L, Pascual V. The innate immune system in SLE: type I interferons and dendritic cells. *Lupus* 2008;17:394-9.
4. Malmstrom V, Trollmo C, Klareskog L. The additive role of innate and adaptive immunity in the development of arthritis. *Am J Med Sci* 2004;327:196-201.
5. Bowes J, Barton A. Recent advances in the genetics of RA susceptibility. *Rheumatology (Oxford)* 2008;47:399-402.
6. Rhodes B, Vyse TJ. The genetics of SLE: an update in the light of genome-wide association studies. *Rheumatology (Oxford)* 2008;47:1603-11.
7. Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM, et al. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. *Nat Genet* 2008;40:1059-61.
8. Musone SL, Taylor KE, Lu TT, Nititham J, Ferreira RC, Ortmann W, et al. Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. *Nat Genet* 2008;40:1062-4.
9. Fernandez A, Quintana G, Rondon F, Restrepo JF, Sanchez A, Matteson EL, et al. Lupus arthropathy: a case series of patients with rhus. *Clin Rheumatol* 2006;25:164-7.
10. Gregersen PK, Behrens TW. Genetics of autoimmune diseases: disorders of immune homeostasis. *Nat Rev Genet* 2006;7:917-28.

11. Becker KG. The common variants/multiple disease hypothesis of common complex genetic disorders. *Med Hypotheses* 2004;62:309-17.
12. Lin JP, Cash JM, Doyle SZ, Peden S, Kanik K, Amos CI, et al. Familial clustering of rheumatoid arthritis with other autoimmune diseases. *Hum Genet* 1998;103:475-82.
13. Payami H, Joe S, Thomson G. Autoimmune thyroid disease in type I diabetic families. *Genet Epidemiol* 1989;6:137-41.
14. Fernando MM, Stevens CR, Walsh EC, De Jager PL, Goyette P, Plenge RM, et al. Defining the role of the MHC in autoimmunity: a review and pooled analysis. *PLoS Genet* 2008;4:e1000024.
15. Becker KG, Simon RM, Bailey-Wilson JE, Freidlin B, Biddison WE, McFarland HF, et al. Clustering of non-major histocompatibility complex susceptibility candidate loci in human autoimmune diseases. *Proc Natl Acad Sci U S A* 1998;95:9979-84.
16. Griffiths MM, Encinas JA, Remmers EF, Kuchroo VK, Wilder RL. Mapping autoimmunity genes. *Curr Opin Immunol* 1999;11:689-700.
17. Dieguez-Gonzalez R, Calaza M, Perez-Pampin E, de la Serna AR, Fernandez-Gutierrez B, Castaneda S, et al. Association of interferon regulatory factor 5 haplotypes, similar to that found in systemic lupus erythematosus, in a large subgroup of patients with rheumatoid arthritis. *Arthritis Rheum* 2008;58:1264-74.
18. Sigurdsson S, Padyukov L, Kurreeman FA, Liljedahl U, Wiman AC, Alfredsson L, et al. Association of a haplotype in the promoter region of the interferon regulatory factor 5 gene with rheumatoid arthritis. *Arthritis Rheum* 2007;56:2202-10.
19. Suarez-Gestal M, Calaza M, Endreffy E, Pullmann R, Ordi-Ros J, Domenico Sebastiani G, et al. Replication of recently identified systemic lupus erythematosus genetic associations: a case control study. *Arthritis Res Ther* 2009;11:R69.
20. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315-24.
21. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263-5.
22. Dupont WD, Plummer WD Jr. Power and sample size calculations: a review and computer program. *Control Clin Trials* 1990;11:116-28.
23. Raman K, Mohan C. Genetic underpinnings of autoimmunity: lessons from studies in arthritis, diabetes, lupus and multiple sclerosis. *Curr Opin Immunol* 2003;15:651-9.
24. Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet* 2001;29:306-9.
25. Todd JA. Statistical false positive or true disease pathway? *Nat Genet* 2006;38:731-3.
26. International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEN), Harley JB, Alarcon-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXX, KIAA1542 and other loci. *Nat Genet* 2008;40:204-10.
27. Kozyrev SV, Abelson AK, Wojcik J, Zaghloul A, Linga Reddy MV, Sanchez E, et al. Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. *Nat Genet* 2008;40:211-6.
28. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, et al. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. *N Engl J Med* 2008;358:900-9.
29. Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Guiducci C, Burtt NP, et al. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. *Nat Genet* 2008;40:1216-23.
30. Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller

- J, et al. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. *Nat Genet* 2007;39:1477–82.
31. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661–78.
32. Sigurdsson S, Nordmark G, Goring HH, Lindroos K, Wiman AC, Sturfelt G, et al. Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. *Am J Hum Genet* 2005;76:528–37.
33. Sawalha AH, Webb R, Han S, Kelly JA, Kaufman KM, Kimberly RP, et al. Common variants within MEC2 confer risk of systemic lupus erythematosus. *PLoS ONE* 2008;3:e1727.
34. Cunninghame Graham DS, Vyse TJ, Fortin PR, Montpetit A, Cai YC, Lim S, et al. Association of LY9 in UK and Canadian SLE families. *Genes Immun* 2008;9:93–102.
35. Cunninghame Graham DS, Akil M, Vyse TJ. Association of polymorphisms across the tyrosine kinase gene, TYK2 in UK SLE families. *Rheumatology (Oxford)* 2007;46:927–30.
36. Plenge RM, Padyukov L, Remmers EF, Purcell S, Lee AT, Karlson EW, et al. Replication of putative candidate-gene associations with rheumatoid arthritis in >4,000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4. *Am J Hum Genet* 2005;77:1044–60.
37. Orozco G, Abelson AK, Gonzalez-Gay MA, Balsa A, Pascual-Salcedo D, Garcia A, et al. Study of functional variants of the BANK1 gene in rheumatoid arthritis. *Arthritis Rheum* 2009;60:372–9.
38. Lee HS, Remmers EF, Le JM, Kastner DL, Bae SC, Gregersen PK. Association of STAT4 with rheumatoid arthritis in the Korean population. *Mol Med* 2007;13:455–60.
39. Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. *N Engl J Med* 2007;357:977–86.
40. Barton A, Thomson W, Ke X, Eyre S, Hinks A, Bowes J, et al. Re-evaluation of putative rheumatoid arthritis susceptibility genes in the post-genome wide association study era and hypothesis of a key pathway underlying susceptibility. *Hum Mol Genet* 2008;17:2274–9.
41. Kobayashi S, Ikari K, Kaneko H, Kochi Y, Yamamoto K, Shimane K, et al. Association of STAT4 with susceptibility to rheumatoid arthritis and systemic lupus erythematosus in the Japanese population. *Arthritis Rheum* 2008;58:1940–6.
42. Orozco G, Alizadeh BZ, Delgado-Vega AM, Gonzalez-Gay MA, Balsa A, Pascual-Salcedo D, et al. Association of STAT4 with rheumatoid arthritis: a replication study in three European populations. *Arthritis Rheum* 2008;58:1974–80.
43. Martinez A, Varade J, Marquez A, Cenit MC, Espino L, Perdigones N, et al. Association of the STAT4 gene with increased susceptibility for some immune-mediated diseases. *Arthritis Rheum* 2008;58:2598–602.
44. Zervou MI, Sidiropoulos P, Petraki E, Vazgiourakis V, Krasoudaki E, Raptopoulou A, et al. Association of a TRAF1 and a STAT4 gene polymorphism with increased risk for rheumatoid arthritis in a genetically homogeneous population. *Hum Immunol* 2008;69:567–71.
45. Palomino-Morales RJ, Rojas-Villarraga A, Gonzalez CI, Ramirez G, Anaya JM, Martin J. STAT4 but not TRAF1/C5 variants influence the risk of developing rheumatoid arthritis and systemic lupus erythematosus in Colombians. *Genes Immun* 2008;9:379–82.
46. Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B, et al. TRAF1-C5 as a risk locus for rheumatoid arthritis: a genomewide study. *N Engl J Med* 2007;357:1199–209.

Follow-up case-control study of the *TNFAIP3* association with Rheumatoid Arthritis: possible role of anti-nuclear antibody status and differences with other diseases

Marian Suarez-Gestal¹, Manuel Calaza¹, Rebeca Dieguez-Gonzalez¹, Gabriel Herrero-Beaumont², Juan D Cañete³, Ana María Ortiz⁴, Jose Luis Marenco⁵, Rafael Caliz⁶, Benjamin Fernandez-Gutierrez⁷, Javier Narvaez⁸, Arturo R de la Serna⁹, Luis Carreño¹⁰, Alejandro Balsa¹¹, Francisco J Blanco¹², Eva Perez-Pampin¹, Beatriz Joven¹³, Federico Navarro¹⁴, Juan J Gomez-Reino^{1,15} and Antonio Gonzalez¹

1 Laboratorio de Investigacion 10 and Rheumatology Unit. Instituto Investigacion Sanitaria-Hospital Clinico Universitario de Santiago. Xantiago de Compostela. Spain. mariansg82@hotmail.com; mc calaza@gmail.com; rebecadg@usc.es;

evappampin@gmail.com

2 Rheumatology Unit. Fundacion Jimenez Diaz. Madrid. Spain. g herrero@fjd.es

3 Rheumatology Unit Hospital Clinic. Barcelona. Spain. JCANETE@clinic.ub.es

4 Rheumatology Unit. Hospital Universitario de la Princesa. Madrid. Spain.

lanult@yahoo.es

5 Rheumatology Unit. Hospital Universitario de Valme. Sevilla. Spain.

josel.marenco.sspa@juntadeandalucia.es

6 Rheumatology Unit. Hospital Universitario Virgen de las Nieves. Granada. Spain.

rcaliz@fundacionhvn.org

7 Rheumatology Unit. Hospital Clinico San Carlos. Madrid. Spain.

bfernandez.hcsc@salud.madrid.org

8 Rheumatology Unit. Hospital Universitario de Bellvitge. Barcelona. Spain.

31577edd@comb.es

9 Rheumatology Unit. Hospital Santa Creu e San Pau. Barcelona. Spain.

arturojj@hotmail.com

10 Rheumatology Unit. Hospital Universitario Gregorio Marañón. Madrid. Spain.

lcarreno.hgugm@salud.madrid.org

11 Rheumatology Unit. Hospital La Paz . Madrid. Spain. al.balsa@telefonica.net

12 Laboratorio de Investigación Osteoarticular y del Envejecimiento. INIBIC-Complejo Hospitalario Universitario A Coruña. Spain. Francisco_Blanco@canalejo.org

13 Rheumatology Unit. Hospital 12 de Octubre. Madrid. Spain.

beatrizjoven@hotmail.com

14 Rheumatology Unit. Hospital Universitario Virgen Macarena. Sevilla. Spain.

federiconavarro@supercable.es

15 Department of Medicine. University of Santiago. Santiago de Compostela. Spain.

Juan.Jesus.Gomez-Reino.Carnota@sergas.es

Corresponding author:

Antonio Gonzalez

Laboratorio de Investigacion 10

Hospital Clinico Universitario de Santiago

Travesía de Choupana sn.

15706-Santiago de Compostela

Spain

Tfn: 34 981 950 903

Fax: 34 981 951 068

Antonio.Gonzalez.Martinez-Pedrayo@sergas.es

anlugon@hotmail.com

ABSTRACT

Introduction: Recent studies have shown that *TNFAIP3* is associated with several autoimmune diseases and pointed to a nsSNP, rs2230926, as the causal polymorphism in SLE, but not in psoriasis. We aimed to complete analysis of this gene in RA.

Methods: Five *TNFAIP3* SNPs, including the nsSNP associated with SLE and the SNP associated with psoriasis, rs610604, and a *PERP* SLE-associated SNP were analyzed in 1657 patients with RA and 1638 controls of Spanish ancestry. Data were complemented with genotypes for other SNPs previously studied in the same samples. Stratification of patients by gender, ACPA, RF, ANA status and by *HLA-DRB1* genotypes was done.

Results: Association was observed with rs610604, but not with SLE associated SNPs or haplotypes. However, association with rs610604 differed from the previously reported in psoriasis in direction of change. Haplotype analysis and genetic models of association indicated that the rs610604 SNP reflects best the association between RA and *TNFAIP3*. ANA status identified differential association of SNPs in the 6q23 region better than ACPA status, although both were correlated and the associations reinforced in these stratified analyses were unrelated with the described in SLE.

Conclusions: The association of *TNFAIP3* with RA was different from the observed in SLE and in psoriasis. Identification of the involved mechanisms will unravel the relationships between these autoimmune diseases. In addition, ANA status was the best discriminator of differential associations in the 6q23 region. Therefore, ANA status could be useful to understand the known genetic heterogeneity of RA patients.

INTRODUCTION

Recent research has uncovered several rheumatoid arthritis (RA) genetic factors that are consistently replicated in successive studies [1]. This reproducibility allows for experiments aimed to discover molecular mechanisms and therapeutic and clinical implications. One of these confirmed RA susceptibility factors is particularly intriguing because it has been identified in a region of chromosome 6q23 devoid of genes and because it includes variants showing two independent effects [2-4]. These two effects are identified by the rs6920220 SNP, which minor allele is associated with increased RA susceptibility, and by the rs10499194 SNP (or the rs13207033 SNP), which minor allele is associated with RA protection. In addition to these clearly confirmed associations, we and others have found evidence for a combined effect of these two SNPs in the intergenic locus with SNPs in the tumour necrosis factor- α -induced protein 3 (*TNFAIP3*) gene, which is 185 kb downstream [5-6]. *TNFAIP3* is an important protein in the regulation of immune reactions through inhibition of NF κ B signals at multiple levels and, therefore, a good candidate for an important role in RA. In addition, association with the polymorphisms in the 6q23 intergenic region, in *TNFAIP3* or in both has been also observed in systemic lupus erythematosus (SLE) [7-9], psoriasis [10], celiac disease [11], type 1 diabetes [12] and juvenile idiopathic arthritis [13-14], indicating that genetic variation in this broad 6q23 region influences susceptibility to multiple immune-mediated diseases.

We aimed to clarify the complex genetics of RA in this region by adding to our previous analysis [5], which was done with tagSNPs, the most informative SNPs from new reports in other diseases and by analysing specific subphenotypes that could be shared between them. The SLE studies have been the most exhaustive and rewarding because they have found strong association with SNPs at the two loci, the intergenic

region and the *TNFAIP3* gene, and with a third locus: the *PERP* gene that is further 207 Kb downstream from *TNFAIP3* [7-9]. In addition, these SLE studies showed strong association with a *TNFAIP3* nsSNP (rs2230926 or its proxy in our samples, rs5029939) [7-9] that was shown to hinder the inhibitory effect of TNFAIP3 on NFkappaB [8]. These results make of this nsSNP the most likely *TNFAIP3* causal variant in SLE susceptibility and suggests a potential mechanism of action. In addition, our new study has included other SNPs because psoriasis shows no association with any of the SNPs that are associated with SLE in the intergenic region or in the *TNFAIP3* locus, but with an unlinked intronic SNP, rs610604 [10]. Consequently, we have included in this follow-up study of RA the most associated SNP in *PERP*, rs6922466, and five *TNFAIP3* SNPs: the nsSNP rs2230926 and two other SNPs strongly associated to SLE rs7749323 and rs5029939, the rs610604 SNP associated to psoriasis and a tagSNP, rs644340, to cover more completely the locus. In addition, we have explored if stratification of patients with RA according to the presence of anti-nuclear antibodies (ANA), and SLE-associated *HLA-DRB1* alleles allowed to identify subgroups of patients with differential association to any of these SNPs. The six SNPs were analyzed in 1657 patients with RA and 1638 controls. Results showed association with the psoriasis susceptibility SNP in *TNFAIP3*, but the direction of change was the opposite. In addition, there was no association with SLE susceptibility SNPs in *TNFAIP3* or in *PERP*. Presence of ANA identified subgroups of patients with differential associations in this region, but these associations did not match with the described in SLE. These results suggest the possibility that the causal variants in this locus may be different in RA, psoriasis and SLE.

MATERIAL AND METHODS

DNA samples: Samples used in this study (1657 patients with RA and 1638 controls) have already been described [5]. All patients met the revised American College of Rheumatology 1987 classification criteria [15]. Participants gave their written informed consent and the respective ethical committees approved the study.

Genotyping: Six SNPs were genotyped in multiplex PCR and single-base extension reactions as described [5]. PCR primers and hybridization probes are detailed in supplementary table 1. HLA-DRB1 was studied by Sequence Based Typing (SBT) using the AlleleSEQR HLA-DRB1 Typing kit (Abbott, Germany), which includes a single PCR amplification for all the alleles of the second exon of DRB1 and bidirectional sequencing. Samples with ambiguous genotypes due to the impossible discrimination between combinations of different pairs of alleles were additionally sequenced with group-specific primers (AlleleSEQR HLA-DRB1 GSSP, Abbott, Germany). Some samples required additional disambiguation based on a large excess in the frequency of one of the allele combinations over the alternative in our Spanish sample.

Statistical analysis: Quality control, comparison of allele and genotype frequency data, conditional logistic regression, haplotype analysis and statistical power estimation were carried out as described [5]. Data from our previous study on this region was included in the conditional and haplotype analyses.

RESULTS

The five *TNFAIP3* SNPs and one *PERP* SNP were genotyped with a 99.9 % call rate and were in accordance with Hardy-Weinberg equilibrium in controls. Comparison of allele frequencies between the 1657 patients with RA and the 1638 controls showed weak significant differences only in two SNPs: rs610604 and rs644340 (O.R. = 0.89

and 0.90, respectively; Table 1) that are, respectively, the SNP associated with psoriasis and the tagSNP included to complete coverage of the locus. Both showed a modest protective effect of their minor allele. No association was observed with any of the other three *TNFAIP3* SNPs that are the most associated in SLE studies (Table 1). As these three SLE susceptibility SNPs have minor allele frequencies below 5 %, the power of these analyses could be a matter of concern. However, it was possible to exclude with great confidence a similar effect than the reported in SLE (power > 99.9 % to detect an effect of O.R. = 2.0 with $p < 0.001$). This is consequence of the large size of our study and the considerable magnitude of the effects reported in SLE (O.R. ≥ 2.0)[7-9]. In addition, the *PERP* rs6922466 SNP did not show allele differences between patients with RA and controls in our study. Again, power of our study was enough (> 90 % for $p < 0.001$) to detect an effect as large as the reported in SLE (O.R. = 0.77)[8]. Genotype analyses yielded similar results (not shown).

Integration of these and our previous results showed that rs610604 was the *TNFAIP3* SNP most strongly associated with RA. However, it was in tight LD with other two associated SNPs, rs582757 and rs644340, and their individual effects were redundant by conditional logistic regression analysis (not shown). That is, association with any of the three SNPs disappeared when any of the other two was included in a multilocus model. These results suggest that there is a single causal polymorphism in *TNFAIP3* for the RA association.

Analysis of *TNFAIP3* haplotypes with the ten SNPs (five from our previous report, and five from this study) showed no difference between cases and controls (Table 2). In contrast, the SLE associated haplotype defined by the rare allele of rs7749323 was not associated with RA. We also combined the enlarged data from the *TNFAIP3* locus with our results from the 6q23 intergenic locus [5]. As could be expected by the redundancy

between rs582757, rs610604 and rs644340, any of these three SNPs was present in the best regression logistic models together with SNPs of the intergenic locus (not shown). These models showed association with RA in both the 6q23 intergenic region and in the *TNFAIP3* locus through epistasis or interdependence of their effects. No contribution from the newly studied SNPs in *TNFAIP3* or in *PERP* was found. Therefore, none of the analyses performed, single SNP, haplotype or multilocus, have provided new independent associations or new interactions over the already described.

Given the discrepant associations between SLE and RA in the *TNFAIP3* locus we were specially interested in analyzing RA subphenotypes, with particular interest in patients with SLE-like characteristics (clinical features are shown in supplementary table 2). In our previous analyses we stratified by gender, ACPA and RF status and we detected associations between women and ACPA positive patients and SNPs of the intergenic locus, but none with SNPs of the *TNFAIP3* locus. Similarly, none of the new SNPs was differentially associated when stratified by these three clinical variables (not shown). In addition, we have analyzed all the SNPs after stratification by the presence of ANA, the shared epitope and SLE-associated *HLA-DRB1* alleles (*0301, *1501 and *0801).

Stratification by ANA status identified markedly different subgroups of patients regarding association with the intergenic and the *TNFAIP3* loci. There were 5 SNPs with different frequencies between ANA positive and ANA negative patients (Table 3). In addition, 4 SNPs were different between ANA positive patients and controls, and 2 between ANA negative patients and controls. Surprisingly, none of the three *TNFAIP3* SNPs associated with SLE, rs2230926, rs7749323 and rs5029939, were significantly influenced by stratification according to ANA status. Conditional analysis showed that the 4 SNPs different in ANA positive patients were explained by rs13207033, which showed the most marked difference with controls. This SNP was associated only with

the RA patients that were ANA positive in our collection of samples (Table 3). On the other side, the two SNPs that were specially associated with the ANA negative subgroup of patients were redundant, although rs610604 explained better the association in conditional analysis than rs11970361 (not shown). This SNP, which explained all association in the *TNFAIP3* locus, was only associated in the ANA negative subgroup, with no difference between the ANA positive patients and the controls (Table 3). None of the stratifications on HLA alleles identified any differential association in the two loci (not shown).

The marked effect of ANA status motivated us to test if the ANA positive patients were different from the ANA negative ones in the presence of ACPA, RF, SE alleles and SLE associated HLA-DRB1 alleles (Table 4). The comparisons showed that ANA positive patients were similar to the ANA negative patients in all these features except in the presence of RF that was more common among the ANA positive subjects. Therefore, there was not any indication to suspect of the classification as RA of the ANA positive patients.

Our previous analyses have showed that association of rs13207033 was restricted to ACPA positive patients. Therefore, we checked how the two subgroups of antibody positive patients were related. Most ANA positive patients were also ACPA positive (66.2 %, Table 4), but association was not different in the ANA positive patients that were ACPA negative and in the positive for the two autoantibodies (O.R. = 0.67 vs 0.74, respectively). Therefore, the ANA status had a more marked effect in association with this SNP than the ACPA status and was able to explain it entirely.

DISCUSION

Our follow-up study has not identified any new independent association with RA in the *TNFAIP3* locus or in the *PERP* gene. This indicates that although association in this

region of chromosome 6q23 is complex, it could be tractable. In addition, our study has helped to clarify it by showing that the pattern of association with RA in *TNFAIP3* does not resemble SLE or psoriasis. The most associated SNP was rs610604, as described in psoriasis, but the risk alleles were the opposite. The rare allele has been described as associated with increased risk in psoriasis, but it was associated with protection in RA. In addition, the SLE susceptibility SNPs were not associated with RA. This lack of association of the SLE SNPs was observed in spite of a very good power for a similar effect. Also psoriasis has been found to lack association with the SLE SNPs in *TNFAIP3*. This pattern is important because suggests that association of the *TNFAIP3* locus with several diseases cannot be assumed to be due to the same polymorphism and mechanisms. However, these results require replication in independent studies given the weak association in the locus we have found and lack of concordance with the results of two recent studies that showed association of RA with the SLE-associated rs2230926, weak in a UK study ($P = 0.03$, O.R. = 1.21)[6] and stronger in the Japanese ($P = 2.6 \times 10^{-5}$, O.R. = 1.35)[16].

The type of comparative analysis between different autoimmune diseases is important because these diseases share multiple characteristics, but they also have distinctive features and their identification can be very useful to understand the possibilities and limitations of transferring knowledge and treatments. The commonality of genetic factors between autoimmune diseases is supported by multiple shared associations. In the case of RA and SLE, there is clear evidence of a set of loci that are shared. They include *PTPN22*, *STAT4*, the 6q23 intergenic locus, *C8orf13-BLK* and *PXK*, and, probably, *IRF5* and *CTLA4* [17-18]. Other associations seem to be specific either of SLE, as the *ITGAM* or *IRAK1-MECP2* loci, or of RA, as the *C5-TRAF1* locus. Association with *HLA-DRB1* is peculiar as it involves different alleles in RA and in

SLE. As mentioned, our results suggest that association with *TNFAIP3* will be of this type too, present in the two diseases but with different alleles. The genetic relationship between RA and psoriasis is less clear. They do not share any of the clearly demonstrated susceptibility loci, not even *PTPN22*. An apparent exception could have been association with the *TNFAIP3* locus (psoriasis does not show association with SNPs in the 6q23 intergenic locus) [10], but the same variant showed opposed effects in our RA patients than the described in psoriasis.

Heterogeneity in the association with the 6q23 region has also been observed within the patients with RA since the first reports [2-6, 16]. It seems that one of the two independent associations in the intergenic 6q23 locus, identified by rs13207033 or rs10499194, is predominant in patients with ACPA [3, 5, 16]. Now, we have found that it was also more marked in patients with ANA. In fact, the presence of ANA was able to account for the effect of ACPA status in this association in spite of identifying a much smaller fraction of the patients. In addition, ANA status was differentially associated with SNPs in the *TNFAIP3* locus, but not with the SLE-associated SNPs in this locus, as could have been expected. These results are intriguing because they suggest that the presence of ANA is a possible marker of a subgroup of RA patients and because they add to the evidence of heterogeneous effects of genetic variation in the 6q23 region. However, these results should be interpreted with caution because not all reports are concordant in the relationship between ACPA subphenotypes and SNPs in the region [6], and because there were limitations in our stratification on ANA. These limitations include the performance of multiple comparisons between subgroups of patients, lack of ANA data for a fraction of the patients and assessment of ANA status at different times in the evolution of RA. Specifically, some of the ANA data were obtained after treatment with TNF antagonists, which is known to induce production of these

autoantibodies [19]. On the contrary, we do not find reasons to doubt of the classification of the ANA positive patients as genuine RA: apart from being classified as RA on clinical grounds their frequencies of ACPA, RF and SE were typical of RA.

CONCLUSIONS

In summary, this follow-up study indicates that only three association signals in the 6q23 region contribute to RA susceptibility: two in the intergenic locus and one in *TNFAIP3*, as previously reported [6, 17]. Participation of the SLE-associated SNP in *PERP* was excluded. Results also suggest that the association in *TNFAIP3* is different from the described in SLE and in psoriasis, although this result will require confirmation. Our study also highlights the possibility that ANA status could be useful in stratifying RA patients because it identified differential patterns of association in the two 6q23 loci.

ABBREVIATIONS

TNFAIP3: tumour necrosis factor- α -induced protein 3

RA: rheumatoid arthritis

SLE: systemic lupus erythematosus

OR: odds ratio

CI: confidence interval

MAF: minor allele frequency

COMPETING INTEREST

The author(s) declare that they have no competing interests.

AUTHORS' CONTRIBUTION

MS-G participated in design of the study, genotyped the samples, participated in the interpretation of the results and in writing the manuscript. MC participated in the statistical analysis and in the interpretation of results. RD-G performed HLA-DRB1 genotyping. GH-B, J-DC, A-MO, J-LM, RC, BF-G, JN, AR-S, LC, AB, F-JB, EP-P, J-LP and FN participated in the acquisition of clinical data and collection of samples and in the analysis and interpretation of results. JJG-R coordinated the acquisition of clinical data and collection of samples and participated in the analysis and interpretation of results. AG participated in the design of the study and in the coordination of acquisition of clinical data and collection of samples, supervised genotyping, statistical analysis, interpretation of results and writing of the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We thank Carmen Pena-Pena for her excellent technical assistance. M.S-G is the recipient of a FPU pre-doctoral bursary of the Spanish Ministry of Education; M.C. is the recipient of an "Isabel Barreto" bursary of the Government of Galicia. This project was supported by grants PI080744 and by RETICS Program, RD08/0075 (RIER) from the Instituto de Salud Carlos III (Spain) with participation of funds from FEDER (European Union).

REFERENCES

1. Bowes J, Barton A: **Recent advances in the genetics of RA susceptibility.** *Rheumatology (Oxford)* 2008, **47**(4):399-402.
2. **Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls.** *Nature* 2007, **447**(7145):661-678.
3. Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller J, Pe'er I, Burtt NP, Blumenstiel B, DeFelice M, Parkin M, Barry R, Winslow W, Healy C,

- Graham RR, Neale BM, Izmailova E, Roubenoff R, Parker AN, Glass R, Karlson EW, Maher N, Hafler DA, Lee DM, Seldin MF, Remmers EF, Lee AT, Padyukov L, Alfredsson L, Coby J *et al*: **Two independent alleles at 6q23 associated with risk of rheumatoid arthritis**. *Nat Genet* 2007, **39**(12):1477-1482.
4. Thomson W, Barton A, Ke X, Eyre S, Hinks A, Bowes J, Donn R, Symmons D, Hider S, Bruce IN, Wilson AG, Marinou I, Morgan A, Emery P, Carter A, Steer S, Hocking L, Reid DM, Wordsworth P, Harrison P, Strachan D, Worthington J: **Rheumatoid arthritis association at 6q23**. *Nat Genet* 2007, **39**(12):1431-1433.
5. Dieguez-Gonzalez R, Calaza M, Perez-Pampin E, Balsa A, Blanco FJ, Canete JD, Caliz R, Carreno L, de la Serna AR, Fernandez-Gutierrez B, Ortiz AM, Herrero-Beaumont G, Pablos JL, Narvaez J, Navarro F, Marengo JL, Gomez-Reino JJ, Gonzalez A: **Analysis of TNFAIP3, a feedback inhibitor of nuclear factor-kappaB and the neighbor intergenic 6q23 region in rheumatoid arthritis susceptibility**. *Arthritis Res Ther* 2009, **11**(2):R42.
6. Orozco G, Hinks A, Eyre S, Ke X, Gibbons LJ, Bowes J, Flynn E, Martin P, Wilson AG, Bax DE, Morgan AW, Emery P, Steer S, Hocking L, Reid DM, Wordsworth P, Harrison P, Thomson W, Barton A, Worthington J: **Combined effects of three independent SNPs greatly increase the risk estimate for RA at 6q23**. *Hum Mol Genet* 2009, **18**(14):2693-2699.
7. Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM, Burt NP, Guiducci C, Parkin M, Gates C, Plenge RM, Behrens TW, Wither JE, Rioux JD, Fortin PR, Graham DC, Wong AK, Vyse TJ, Daly MJ, Altshuler D, Moser KL, Gaffney PM: **Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus**. *Nat Genet* 2008, **40**(9):1059-1061.
8. Musone SL, Taylor KE, Lu TT, Nititham J, Ferreira RC, Ortmann W, Shifrin N, Petri MA, Kamboh MI, Manzi S, Seldin MF, Gregersen PK, Behrens TW, Ma A, Kwok PY, Criswell LA: **Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus**. *Nat Genet* 2008, **40**(9):1062-1064.
9. Bates JS, Lessard CJ, Leon JM, Nguyen T, Battist L, Rodgers J, Kaufman KM, James JA, Gilkeson GS, Kelly JA, Humphrey MB, Harley JB, Gray-McGuire C, Moser KL, Gaffney PM: **Meta-analysis and imputation identifies a 109 kb risk haplotype spanning TNFAIP3 associated with lupus nephritis and hematologic manifestations**. *Genes Immun* 2009, **10**(5):470-477.
10. Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, Goldgar D, Gudjonsson JE, Li Y, Tejasvi T, Feng BJ, Ruether A, Schreiber S, Weichenthal M, Gladman D, Rahman P, Schrod SJ, Prahalad S, Guthery SL, Fischer J, Liao W, Kwok PY, Menter A, Lathrop GM, Wise CA, Begovich AB, Voorhees JJ, Elder JT, Krueger GG, Bowcock AM, Abecasis GR: **Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways**. *Nat Genet* 2009, **41**(2):199-204.
11. Trynka G, Zhernakova A, Romanos J, Franke L, Hunt KA, Turner G, Bruinenberg M, Heap GA, Platteel M, Ryan AW, de Kovel C, Holmes GK, Howdle PD, Walters JR, Sanders DS, Mulder CJ, Mearin ML, Verbeek WH, Trimble V, Stevens FM, Kelleher D, Barisani D, Bardella MT, McManus R, van Heel DA, Wijmenga C: **Celiac disease-associated risk variants in TNFAIP3 and REL implicate altered NF-kappaB signalling**. *Gut* 2009, **58**(8):1078-1083.

12. Fung EY, Smyth DJ, Howson JM, Cooper JD, Walker NM, Stevens H, Wicker LS, Todd JA: **Analysis of 17 autoimmune disease-associated variants in type 1 diabetes identifies 6q23/TNFAIP3 as a susceptibility locus.** *Genes Immun* 2009, **10**(2):188-191.
13. Prahalad S, Hansen S, Whiting A, Guthery SL, Clifford B, McNally B, Zeff AS, Bohnsack JF, Jorde LB: **Variants in TNFAIP3, STAT4, and C12orf30 loci associated with multiple autoimmune diseases are also associated with juvenile idiopathic arthritis.** *Arthritis Rheum* 2009, **60**(7):2124-2130.
14. Hinks A, Eyre S, Ke X, Barton A, Martin P, Flynn E, Packham J, Worthington J, Thomson W: **Overlap of disease susceptibility loci for rheumatoid arthritis (RA) and juvenile idiopathic arthritis (JIA).** *Ann Rheum Dis* 2009.
15. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, Healey LA, Kaplan SR, Liang MH, Luthra HS, et al.: **The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis.** *Arthritis Rheum* 1988, **31**(3):315-324.
16. Shimane K, Kochi Y, Horita T, Ikari K, Amano H, Hirakata M, Okamoto A, Yamada R, Myouzen K, Suzuki A, Kubo M, Atsumi T, Koike T, Takasaki Y, Momohara S, Yamanaka H, Nakamura Y, Yamamoto K: **The association of a nonsynonymous single-nucleotide polymorphism in TNFAIP3 with systemic lupus erythematosus and rheumatoid arthritis in the Japanese population.** *Arthritis Rheum* 2010, **62**(2):574-579.
17. Suarez-Gestal M, Calaza M, Dieguez-Gonzalez R, Perez-Pampin E, Pablos JL, Navarro F, Narvaez J, Marengo JL, Herrero-Beaumont G, Fernandez-Gutierrez B, Lamas JR, de la Serna AR, Ortiz AM, Carreno L, Canete JD, Caliz R, Blanco FJ, Balsa A, Gomez-Reino JJ, Gonzalez A: **Rheumatoid arthritis does not share most of the newly identified systemic lupus erythematosus genetic factors.** *Arthritis Rheum* 2009, **60**(9):2558-2564.
18. Gregersen PK, Amos CI, Lee AT, Lu Y, Remmers EF, Kastner DL, Seldin MF, Criswell LA, Plenge RM, Holers VM, Mikuls TR, Sokka T, Moreland LW, Bridges SL, Jr., Xie G, Begovich AB, Siminovitch KA: **REL, encoding a member of the NF-kappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis.** *Nat Genet* 2009, **41**(7):820-823.
19. Charles PJ, Smeenk RJ, De Jong J, Feldmann M, Maini RN: **Assessment of antibodies to double-stranded DNA induced in rheumatoid arthritis patients following treatment with infliximab, a monoclonal antibody to tumor necrosis factor alpha: findings in open-label and randomized placebo-controlled trials.** *Arthritis Rheum* 2000, **43**(11):2383-2390.

Table 1. Comparison of the minor allele frequencies of the five *TNFAIP3* and the *PERP* SNPs between RA cases and controls

	MAF % (n/N) ^a		O.R. (95% C.I.) ^b	<i>P</i>
SNP	RA patients	Controls		
<i>TNFAIP3</i>				
rs2230926	4.9 (161/3312)	4.5 (147/3274)	1.09 (0.9-1.4)	ns ^c
rs5029939	4.8 (158/3308)	4.5 (147/3274)	1.07 (0.8-1.3)	ns
rs610604	31.6 (1045/3308)	34.3 (1122/3272)	0.89 (0.8-1.0)	0.020
rs644340	35.7 (1184/3314)	38.3 (1253/3272)	0.90 (0.8-1.0)	0.031
rs7749323	2.7 (89/3314)	2.4 (77/3274)	1.15 (0.8-1.6)	ns
<i>PERP</i>				
rs6922466	26.3 (871/3310)	27.7 (906/3274)	0.93 (0.8-1.0)	ns

^a MAF = Minor Allele Frequency, n = number of minor alleles, N = total number of alleles

^b O.R. = Odds Ratio, C.I. = Confidence Interval

^c ns = not significant

Table 2. Comparison of the *TNFAIP3* haplotype frequencies between patients with RA and controls

Alleles at each SNP										Controls N (%)	RA patients N (%)
1	2	3	4	5	6	7	8	9	10		
T	C	G	C	T	T	A	A	G	G	1150 (35.1%)	1250 (37.7%)
C	C	G	C	T	T	A	A	G	G	843 (25.7%)	837 (25.3%)
T	C	G	C	T	C	G	G	G	G	453 (13.8%)	417 (12.6%)
T	C	G	C	T	C	G	G	A	G	421 (12.9%)	391 (11.8%)
T	T	C	C	T	T	G	G	G	G	178 (5.4%)	166 (5.0%)
T	C	C	G	G	T	A	G	G	A	72 (2.2%)	88 (2.7%)
T	C	C	G	G	T	G	G	G	G	46 (1.4%)	41 (1.2%)
T	C	C	G	G	T	A	G	G	G	23 (0.7%)	29 (0.9%)
Other haplotypes										90 (2.7%)	95 (2.9%)
Total										3276 (100%)	3314 (100%)

^a Corresponding SNPs: 1: rs600144; 2: rs11970361; 3: rs11970411; 4: rs5029939; 5: rs2230926; 6: rs582757 ; 7: rs610604 ; 8: rs644340 ; 9: rs17780429; 10: rs7749323

Table 3. Analysis of the allele frequencies in the 6q23 intergenic and *TNFAIP3* loci after stratification of patients with RA by ANA status

SNP	MAF % (n/N) ^a			ANA ⁺ vs. controls		ANA ⁻ vs. controls		
	ANA ⁺	ANA ⁻	Controls	O.R. (95% C.I.) ^b	P	O.R. (95% C.I.)	P	P ANA ⁺ vs ANA ⁻
rs13207033	23.0 (100/434)	28.2 (305/1082)	29.7 (914/3082)	0.71 (0.6-0.9)	0.0044	0.93 (0.8-1.1)	0.36	0.041
rs6920220	24.4 (110/450)	21.2 (240/1134)	20.1 (639/3184)	1.29 (1.0-1.6)	0.032	1.07 (0.9-1.3)	0.43	0.16
rs694069	33.4 (149/446)	38.7 (436/1126)	40.0 (1276/3188)	0.75 (0.6-0.9)	0.0073	0.95 (0.8-1.1)	0.44	0.05
rs600144	20.9 (95/454)	27.7 (329/1188)	26.8 (857/3202)	0.72 (0.6-0.9)	0.008	1.05 (0.9-1.2)	0.54	0.0051
rs11970361	8.3 (38/458)	4.3 (51/1192)	6.0 (194/3242)	1.42 (1.0-2.0)	0.056	0.70 (0.5-1.0)	0.028	0.0012
rs610604	35.8 (164/458)	30.3 (361/1190)	34.3 (1122/3272)	1.07 (0.9-1.3)	0.52	0.83 (0.7-1.0)	0.013	0.033

^a MAF = Minor Allele Frequency, n = number of minor alleles, N = total number of alleles^b O.R. = Odds Ratio, C.I. = Confidence Interval

Table 4. Comparison of the frequency of RA or SLE distinctive characteristics between ANA positive and negative patients

	ANA % (n/N) ^a		<i>P</i>
	positive	negative	
ACPA +	66.2 (151/228)	59.5 (353/593)	0.08
RF +	78.7 (166/211)	63.0 (349/554)	0.00003
SE +	44.4 (40/90)	54.2 (212/391)	0.09
HLA-DRB1 *0301.*1501.*0801	27.0 (24/89)	26.8 (110/410)	0.9

^a n = number of subjects with the characteristic; N = total number of available subjects with complete information

Additional files provided with this submission:

Additional file 1: Supplementary Tables.doc, 44K

<http://arthritis-research.com/imedia/1742438513355044/supp1.doc>

Supplementary Table 1. Primers and probes used for genotyping the *TNFAIP3* and *PERP* SNPs represented in the 5' to 3' direction

Locus	SNP	PCR primers		Minisequencing Probe
		Forward Primer	Reverse Primer	
<i>TNFAIP3</i>	rs5029939	ccctgtgtgctcctccttag	tgaatgacacccaactgcaaaag	TTACCTATGAAcaccacactgcaaaaggagcccatgggaccttggttctagcttaa \$
<i>TNFAIP3</i>	rs2230926	ttgctgggtctttacatgcag	cccgtttcaacaaatttcctg	aggacacagacttggtactactgaggaaggcgctgt
<i>TNFAIP3</i>	rs610604	tggatggaccaagattgatt	tgtggattatggcaagaattgta	TTACCTATGATTGATCGTGTGATATCCGattttagaaccattccactgaaatgcaggcctgtgatgaagagct \$
<i>TNFAIP3</i>	rs644340	tagctggaaattacaggcgctc	aggaggtgaggtgcgcag	TTACCTATGATTGATgaggctgaggtgcgcagatcacttaaggcccaagagtcaaaaac \$
<i>TNFAIP3</i>	rs7748323	gtctgagccttgcctcgag	tgggtctcctctctctgtgct	ggtagaccatggaaatccacattaaagaca
<i>PERP</i>	rs6922466	gtctgttcgcgcaccattac	acagctgaattcaaaaagtggg	ttaatgtatcattcagtttttgtctctggggaagtttttggttttggttt

Primers were selected with Primer3 and Oligos. They were checked to avoid formation of dimers between the primers included in the reaction.
\$ These oligonucleotides were extended with a 5' tail that has not homology with human sequences (these tails are in capital letters)

Table 2. Clinical characteristics of the patients with RA included in the study

Clinical characteristic (scale)	Value	Available data
Female (%)	75.0	1509
Age of disease onset (median, IQR)	48, 38-58	1460
Morning stiffness (%)	96.0	1441
Arthritis of 3 or more joint areas (%)	99.7	1441
Arthritis of hand joints (%)	99.3	1441
Symmetric arthritis (%)	99.0	1441
Rheumatoid nodules (%)	20.2	1301
Rheumatoid factor (%)	72.7	1546
Erosions (%)	70.3	1530
S. Sicca (%)	8.8	1533
Interstitial pneumonitis (%)	2.8	1533
Shared epitope (carrier %)	54.3	610
ACPA (%)	66.3	671
ANA (%)	27.8	825
*0301/*0501/*0801 (carrier %)	27.1	546

DISCUSIÓN

1. Discusión

El LES es una enfermedad de etiología compleja en la que el componente genético juega un papel importante, como se ha visto en estudios de agregación familiar y de gemelos (Lawrence JS y col., 1987; Wandstrat A y col., 2001; Deapen D y col., 1992). El *HLA* fue uno de los primeros factores genéticos de susceptibilidad al LES confirmados. Sin embargo, su contribución es menor que en otras enfermedades autoinmunes, como la AR. También se ha demostrado la asociación con LES de *IRF5*, aunque la complejidad de los efectos descubiertos no ha permitido todavía esclarecer su papel en la enfermedad. Otras variantes de susceptibilidad al LES conocidas desde hace años son los factores del complemento *C1q*, *C2* y *C4*, los receptores de baja afinidad para la IgG y el inmunomodulador *PTPN22*. En los últimos años, los avances en la tecnología de genotipado han permitido analizar una gran cantidad de SNPs en un gran número de muestras. En 2008 se han publicado 4 GWAS en LES que han dado a conocer un número considerable de nuevas variantes de susceptibilidad (Hom G y col., 2008, SLEGEN y col., 2008, Kozyrev SV y col., 2008 y Graham RR y col., 2008): *ITGAM*, *KIAA1542*, *PXK*, *1q25.1*, *C8orf13-BLK*, *TNFAIP3* y *BANK1*. A la vez, estudios de asociación de genes candidato han sugerido la implicación en el LES de otros genes, como *TNFSF4* (Cunninghame Graham DS y col., 2008^a), *MECP2-IRAK1* (Jacob CO y col., 2009; Sawalha AH y col., 2008) y *LY9* (Cunninghame Graham DS y col., 2008^b). La consolidación de estas asociaciones necesita de estudios de replicación. Así, uno de los objetivos de este trabajo ha sido confirmar las nuevas señales de asociación en una cohorte independiente de muestras de LES. Se seleccionaron para ello 10 SNPs representativos de 10 factores genéticos (*ITGAM*, *KIAA1542*, *PXK*, *1q25.1*, *C8orf13-BLK*, *STAT4*, *TYK2*, *LY9*, *MECP2* y *BANK1*). Los resultados obtenidos

confirmaron la asociación con el LES de todos estos factores genéticos, excepto *LY9*. Para 6 de ellos (*PXK*, *1q25.1*, *C8orf13-BLK*, *TYK2*, *MECP2* y *BANK1*) esta confirmación fue importante ya que sólo se habían visto asociados en un único estudio o mostraban una asociación inconsistente.

También se trató de confirmar la existencia de una interacción epistática entre *IRF5* y *TYK2* (Hellquist A y col., 2009), ya que se ha sugerido que parte de la heredabilidad no explicada podría deberse a la interacción entre factores genéticos. Para ello se utilizó la comparación de los modelos de regresión logística con y sin el término de interacción y el análisis con el *software* LRASSOC, que también compara modelos con y sin interacción. Los resultados obtenidos no confirmaron la epistasis entre estos dos factores genéticos, indicando su contribución independiente a la susceptibilidad al LES.

PDCD1 fue el primer factor de susceptibilidad al LES identificado a partir de estudios de ligamiento. Sin embargo, las publicaciones posteriores han presentado resultados controvertidos. El polimorfismo más estudiado ha sido el SNP PD1.3, debido a su localización en una región con posible función de *enhancer* (Prokunina L y col., 2002). Con el objetivo de explorar la funcionalidad de PD1.3 se llevaron a cabo ensayos EMSA y de gen *reporter*. Los resultados obtenidos no indicaron un papel funcional para este SNP.

También se buscó una explicación a la heterogeneidad clínica del LES desde dos perspectivas que tienen en cuenta el componente genético. Por un lado, se evaluó la influencia de los SNPs confirmados de susceptibilidad al LES en la prevalencia de 11 variables clínicas: 10 fenotipos definidos por los criterios de clasificación de la ACR (excepto los ANAs) y la edad de inicio de la enfermedad. Se encontraron asociaciones de alelos de susceptibilidad con

variables clínicas, aunque éstas fueron débiles. Por otra parte, se trató de buscar si existía una correlación entre las manifestaciones clínicas y la ascendencia noreuropea o sureuropea de los pacientes. Los resultados confirmaron una mayor frecuencia de fotosensibilidad en el norte y de artritis en el sur, sugiriendo una influencia modesta del fondo genético de estas dos subpoblaciones de europeos en la variabilidad clínica del LES.

Finalmente, se evaluó, mediante un estudio de asociación caso-control, si los nuevos *loci* de susceptibilidad al LES se encontraban asociados con AR, debido a que diversas observaciones sugieren que estas dos enfermedades comparten una fracción de su componente genético. Los resultados mostraron que *STAT4* es el único factor claramente compartido por ambas enfermedades, aunque la combinación de datos con estudios previos también sugerían un efecto modesto de los SNPs de *C8orf13-BLK* y *PXK*. Por otra parte, la región *6q23* también se había identificado en AR y LES (Wellcome Trust Case Control Consortium, 2007; Plenge RM y col., 2007^b; Dieguez-Gonzalez R y col., 2009; Musone SL y col., 2008; Graham RR y col., 2008), aunque las señales de asociación no coincidían exactamente. Un nuevo estudio del *locus TNFAIP3* en AR mostró que las señales de asociación no coincidían con las publicadas en LES.

1.1. Confirmación de factores de susceptibilidad al LES

Cuando un gen se encuentra por primera vez asociado a una enfermedad o su asociación no está clara debido a resultados inconsistentes, son necesarios los estudios de replicación. De esta forma es posible descartar asociaciones debidas al azar o a factores subyacentes no detectados. En este sentido, nuestro objetivo ha sido el de contribuir a la definición de los factores genéticos de

susceptibilidad al LES. Así, se estudiaron los 10 *loci* siguientes: *ITGAM*, *STAT4*, *C8orf13-BLK*, *BANK1*, *PXK*, *KIAA1542*, *1q25.1*, *MECP2*, *LY9* y *TYK2*. La selección de un SNP en cada uno de estos *loci* se hizo en función de su fuerte asociación con LES o su posible papel funcional.

Los resultados de este análisis fueron importantes ya que confirmaron de forma clara la asociación de 9 de los 10 SNPs estudiados, mostrando el mismo alelo de susceptibilidad que el publicado con anterioridad. Los resultados más destacables fueron la confirmación de la asociación de *TYK2*, que hasta el momento ha sido controvertida; la asociación de *MECP2*, del que sólo había un estudio publicado en ese momento; y la asociación de *1q25.1*, *PXK*, *BANK1* y *KIAA1542*, que sólo habían sido identificados en un GWAS. Sin embargo, no se confirmó la asociación de *LY9* con LES, a pesar de que el estudio tenía potencia estadística suficiente para replicar esta señal. Tampoco los GWAS observaron una señal en este *locus*. Los otros tres genes de susceptibilidad estudiados (*ITGAM*, *STAT4* y *C8orf13-BLK*) también fueron replicados y mostraron las ORs más fuertes. Esto explicaría que estas señales hayan sido confirmadas de forma más consistente.

El análisis de estos SNPs se llevó a cabo mediante la aproximación de Mantel-Haenszel, ya que las colecciones de muestras utilizadas proceden de distintos centros europeos. Esta aproximación combina los efectos del SNP en las distintas colecciones de casos y controles, permitiendo que exista heterogeneidad en frecuencias alélicas. La heterogeneidad de efectos entre las colecciones de muestras se evalúa con el análisis de Breslow-Day. Una $P < 0.05$ en este test indicaría que alguna colección tiene un efecto extremo, por lo que tendríamos que excluirla del análisis global. Los resultados obtenidos no

mostraron problemas de heterogeneidad para ninguno de los SNPs, por lo que la agrupación de las distintas colecciones resultó apropiada.

El nsSNP rs2304256 de *TYK2* se localiza en el exón 8 del gen, donde introduce un cambio de valina por fenilalanina en el dominio JH4. Este dominio es crucial no sólo para la interacción de *TYK2* con el receptor de $\text{IFN}\alpha$, *IFNAR1* (Richter MF y col., 1998), sino también para la expresión de *IFNAR1* en las membranas celulares (Ragimbeau J y col., 2003). Sin embargo, todavía no se ha evaluado el posible efecto del nsSNP en estas funciones de *TYK2*. Este SNP fue el más asociado en un estudio de ligamiento y de asociación caso-control de familias escandinavas (Sigurdsson S y col., 2005), pero no se encontró asociado en otro estudio con familias inglesas (Cunninghame Graham DSy col., 2007^a). Sin embargo, en este segundo estudio se encontró asociación con un SNP del intrón 7 de *TYK2* (rs12720270) que, a su vez, no estaba asociado en el estudio de las familias escandinavas (Sigurdsson S y col., 2005). A pesar de que en estos estudios específicos se observan señales de asociación en *TYK2*, éstas no aparecen en los GWAS. Así, el GWAS del SLEGEM excluyó la asociación de rs12720270 (SLEGEM y col., 2008). En este contexto en que se observan discrepancias en las señales de susceptibilidad y ausencia de asociación en los GWAS, es donde la importancia de nuestros resultados se pone de manifiesto al confirmar la asociación del nsSNP. Además, combinando los resultados de los distintos estudios se observó una asociación clara de rs2304256 con LES que pasa el umbral de significación de los GWAS ($P = 2.1 \times 10^{-11}$). Con posterioridad a nuestro trabajo, rs2304256 y rs12720270 también mostraron asociación con LES en un pequeño estudio en finlandeses (Hellquist A y col., 2009) contribuyendo a la confirmación de *TYK2* como factor de susceptibilidad.

La función de la proteína TYK2 está relacionada con las rutas de señalización celular. Así, TYK2 es una quinasa de tirosina de la familia *Janus* que se encuentra unida al dominio citoplasmático de receptores de citocinas, mediando la activación de éstos por sus ligandos. Una vez activada, TYK2 fosforila el receptor al que está unida creando sitios de unión para STATs, que desencadenan la expresión de genes de varias rutas (Yamaoka K y col., 2004). Las rutas en las que se ha visto implicado TYK2 son la del IFN tipo I, IL6, IL10, IL12 e IL23. De modo que la deficiencia de TYK2 en ratones lleva a la alteración de estas rutas y, en consecuencia, a una diferenciación defectuosa de las células Th1 y acelerada de las Th2 (Minegishi Y y col., 2006). Investigaciones futuras indicarán cuál de estas rutas es la responsable de que el alelo de riesgo de *TYK2* aumente la susceptibilidad a LES.

La asociación de *TYK2* fue descubierta en un estudio de genes que participan en la ruta del IFN de tipo I, en el que también se descubrió la asociación de *IRF5* con el LES (Sigursson S y col., 2005). Una publicación reciente observó evidencias de interacción epistática entre *TYK2* e *IRF5* (Hellquist A y col., 2009). Sin embargo, con nuestros datos no se replicó esta interacción entre SNPs de *TYK2* (rs2304256) e *IRF5* (rs10954213, rs10488631, rs729302) asociados al LES. Por tanto, mientras que otros estudios genéticos o funcionales no apoyen la epistasis, no está claro si estos dos factores interaccionan en la misma ruta para conferir una mayor susceptibilidad al LES o si su aportación a la susceptibilidad es por rutas independientes.

Otra contribución importante del presente estudio fue la confirmación de la asociación de *MECP2*. Por una parte, la asociación encontrada en mujeres (OR = 1.26) es muy similar a la publicada también en mujeres de población europeo-americana por Sawalha y col. (OR = 1.29) (Sawalha AH y col., 2008).

Este resultado proporciona una confirmación necesaria ya que no hay señales de asociación con este gen en los GWAS. Otro estudio posterior también ayudó a reforzar este resultado positivo (Webb R y col., 2009). Las dos razones por las que Sawalha y col. consideraron este gen como posible factor de susceptibilidad al LES fueron su localización en el cromosoma X y su participación en la metilación del ADN (Sawalha AH y col., 2008). Esta función podría tener un papel en el LES porque los genes sensibles a metilación sufren una regulación anormal en pacientes con esta enfermedad y la proteína que codifica *MECP2* se une de forma específica a secuencias metiladas del ADN reprimiendo la expresión de los genes que regulan (Klose RJ y col., 2006). Sin embargo, *MECP2* también se une a regiones no metiladas y está asociado con genes que se transcriben de forma activa (Chahrour M y col., 2008). Por tanto, su papel no sería específicamente de silenciamiento génico sino, más bien, de regulación de la transcripción. Por otro lado, Sawalha y col. propusieron que la localización de *MECP2* en el cromosoma X podría explicar el predominio del LES en mujeres (Sawalha AH y col., 2008). No obstante, *MECP2* también está asociado a LES en hombres, como muestran nuestros resultados. La OR es incluso más fuerte en hombres (1.82) que en mujeres (1.26), lo que cuestiona la hipótesis de que este gen contribuye a una mayor prevalencia de LES en mujeres. Esto no sería sorprendente considerando que *MECP2* no se expresa en el cromosoma X inactivado de las mujeres (D'Esposito M y col., 1996), lo que implicaría que los niveles de expresión en ambos sexos serían equivalentes.

No se ha determinado ningún papel funcional para los SNPs asociados en *MECP2* y no está claro si el polimorfismo causal afecta a *MECP2* o a un gen cercano y en LD, llamado *IRAK1*. Este segundo gen podría ser un buen candidato ya que codifica para un mediador clave en las rutas de señalización de los receptores TLR/IL1R. Además, SNPs en *IRAK1* han mostrado asociación

más fuerte que los de *MECP2* ($OR > 1.5$) (Gateva V y col., 2009; Jacob CO y col., 2009). Así pues, es importante que estudios futuros traten de determinar cuál de estos dos genes es el responsable del efecto causal. Después de esto, todavía quedaría identificar el SNP funcional.

Los SNPs rs10798269 en el locus *1q25.1*, rs4963128 en *KIAA1542* y rs6445975 en *PXK* habían sido hallados en un solo GWAS con valores de *P* por debajo de 2×10^{-7} (SLEGEN y col., 2008) y no fueron confirmados en ningún otro artículo. De ahí la importancia de la confirmación observada en nuestro estudio con efectos similares a los publicados ($OR = 0.81$ vs 0.82 , 0.84 vs 0.78 y 1.19 vs 1.25 , respectivamente). Cabe destacar que ninguno de estos 3 polimorfismos se puede relacionar de forma clara con un posible efecto en la patogenia del LES. El locus *1q25.1* se localiza en un intervalo ligado al LES, aunque está lejos de cualquier transcrito conocido. Una relación interesante existe entre el SNP asociado, rs10798269, y otro gen asociado en un estudio previo, *TNFSF4*, ya que sólo se encuentra a 130 kb (Graham DS y col., 2008). De forma que es posible que la señal identificada como correspondiente a *1q25.1* se deba a este último gen. Una situación similar ocurre con *KIAA1542*, que está a unas 20 kb del gen *IRF7* y en LD con él. La asociación del primero puede estar relacionada con la función del segundo como regulador de la expresión de genes de IFN tipo I (SLEGEN y col., 2008). De hecho, recientemente, se ha observado que variantes genéticas de la región *KIAA1542-IRF7*, en combinación con autoanticuerpos característicos del LES, están asociadas con altos niveles de IFN α en suero (Salloum R y col., 2010). Por último, el SNP rs6445975 se localiza en el intrón 5 de *PXK*. Este gen se expresa en una amplia variedad de tejidos, pero poco se sabe de su función. Sólo se conoce que el dominio homólogo a *phox* se une de manera específica a fosfatidilinositol 3-fosfato, induciendo la internalización de proteínas en los

endosomas para su degradación (Takeuchi H y col., 2010). Por tanto, la replicación de estas 3 asociaciones aumenta la necesidad de identificar el polimorfismo causal de cada uno de ellos e investigar sus efectos funcionales.

El SNP rs17266594 en *BANK1* fue otro polimorfismo que se encontró asociado significativamente y que sólo había sido identificado previamente en un único GWAS (Kozyrev SV y col., 2008). El GWAS en el que se descubrió era de baja resolución e incluyó la replicación en una segunda colección de muestras europeas. El SNP rs17266594 fue el más asociado, aunque hay otros 2 posibles SNPs funcionales que también estaban asociados. Nuestros resultados proporcionaron una replicación independiente de dicho SNP, aunque con un efecto ligeramente más modesto (OR = 0.83 vs 0.70). Este SNP podría ser por sí mismo el polimorfismo causal, ya que parece alterar la eficiencia del *splicing* de *BANK1*. Las isoformas alternativas se caracterizan por la pérdida del exón 2 que codifica para el dominio de unión a IP₃R. Pero no ha sido posible discernir entre los 3 SNPs funcionales asociados porque están en fuerte LD.

BANK1 codifica para una proteína estructural de células B con repeticiones de anquirina. Esta proteína estaría implicada en la señalización mediada por el BCR, pero no está claro si ejerce un papel estimulador o inhibidor de la activación de las células B (Yokoyama K y col., 2002; Aiba Y y col., 2006). Puede que su efecto dependa de la interacción con otras proteínas de forma que sean estas interacciones las que determinen el papel o la eficacia de *BANK1* en la ruta de señalización. Una de estas interacciones está detrás de uno de los ejemplos más claros de epistasis identificados en enfermedades complejas y consiste en la interacción epistática entre *BANK1* y *BLK* (Castillejo-Lopez C y col., 2009). Además, esta epistasis está apoyada por

evidencias de interacción física (co-immunoprecipitación y co-localización en un modelo celular *in vitro*) entre las dos proteínas y también con IP₃R.

El único SNP cuya asociación no se confirmó en nuestro estudio fue rs509749 en *LY9*. Este gen se encuentra en *1q23*, uno de los *locus* descritos de forma más consistente en los estudios de ligamiento de LES (y su región sinténica en los estudios de ratones) (Kelly JA y col., 2002). La función de *LY9* en el sistema inmune consiste en mediar la adhesión entre células T, en cuya membrana se expresa, y CPAs. En un estudio basado en familias inglesas y canadienses, el SNP rs509749 mostró la asociación más fuerte de la región *1q23*, pudiendo explicar el ligamiento de ésta con el LES (Cunninghame Graham DS y col., 2008^b). Este SNP es no sinónimo y tiene un impacto predecible en la función de la proteína al alterar su dominio de unión a SAP/SH2D1a, implicada en rutas de señalización de células T. Esta evidencia funcional hacía de este polimorfismo un buen candidato para incluir en nuestro estudio de replicación, a pesar de que el nivel de significación era notablemente más bajo que el encontrado para los otros 9 SNPs estudiados ($P = 0.002$). Esto muestra la dificultad de replicar asociaciones genéticas con valores de P altos.

Los tres factores genéticos restantes (*ITGAM*, *STAT4* y *C8orf13-BLK*) son los que tenían asociaciones más sólidas y que además ya habían sido confirmadas en uno o más estudios previos. El más asociado fue *ITGAM* (OR = 1.70), que se había observado previamente en tres estudios (SLEGEN y col., 2008; Hom G y col., 2008; Nath SK y col., 2008). El SNP que se seleccionó (rs1143679) fue el de menor P en uno de dichos estudios (donde presentó un efecto muy similar al observado aquí, OR = 1.74) (Nath SK y col., 2008). Éste es un nsSNP y se ha hipotetizado que podría alterar la interacción de *ITGAM* con sus ligandos, aunque aún no existe ninguna evidencia de ello. *ITGAM*

codifica para una integrina que se une a una amplia variedad de ligandos (Ehlers MR y col., 2000). Esta molécula media la adhesión leucocitaria, la trasmigración a través del endotelio y la fagocitosis de moléculas opsonizadas con C3bi. Esta última función es importante para la eliminación de células apoptóticas por los macrófagos (Mevorach D y col., 1998), proceso defectuoso en pacientes con LES (Ren Y y col., 2003) que, por tanto, podría contribuir a la patogenia de la enfermedad (Casciola-Rosen LA y col., 1994).

El SNP rs7574865, en el tercer intrón de *STAT4*, mostró la segunda asociación más fuerte con LES (OR = 1.62). Además, es uno de los SNPs de susceptibilidad más claramente establecidos por estudios previos (SLEGEN y col., 2008; Hom G y col., 2008; Remmers EF y col., 2007; Kobayshi S y col., 2008; Palomino-Morales RJ y col., 2008; Sigurdsson S y col., 2008^b; Taylor KE y col., 2008). Este SNP es sólo uno entre varios SNPs altamente correlacionados, que forman un haplotipo asociado con LES (Remmers EF y col., 2008; Sigurdsson S y col., 2008^b; Taylor KE y col., 2008). Sin embargo, no se ha identificado ningún polimorfismo funcional en este *locus*. Sólo se observó la relación de rs7574865 con una mayor expresión de *STAT4* en osteoblastos (Sigurdsson S y col., 2008^b). Por otra parte, se ha mostrado asociación de los mismos SNPs con otras enfermedades autoinmunes, incluyendo AR (Hom G y col., 2008; Remmers EF y col., 2007; Kobayshi S y col., 2008; Lee HS y col., 2007; Martinez A y col., 2008; Orozco G y col., 2008), SS (Korman BD y col., 2008), T1D (Martinez A y col., 2008; Korman BD y col., 2008; Zervou MI y col., 2008^a) y enfermedad inflamatoria intestinal (Martinez A y col., 2008). La relación de *STAT4* con enfermedades autoinmunes también se observó en ratones KO, que son resistentes a múltiples enfermedades autoinmunes, incluyendo AR. Por el contrario, los modelos de LES sufren una enfermedad más grave (Jacob CO y col., 2003; Singh RR y col., 2003). Esta repercusión

opuesta en modelos de distintas enfermedades autoinmunes, podría estar relacionada con la diferenciación de las células T en Th1 o Th2. STAT4 es esencial para la diferenciación de las Th1 (Nishikomori R y col., 2002). Por tanto, este factor de transcripción tendría un impacto diferente en la artritis, enfermedad en la que predominan los efectos de tipo Th1, que en el LES donde podrían predominar los efectos de tipo Th2. Pero esta interpretación necesita revisarse debido al descubrimiento del papel fundamental de los linfocitos Th17 en muchas enfermedades autoinmunes y a la asociación de los mismos alelos de *STAT4* con AR y LES.

Por último, el SNP rs13277113 (OR = 1.34) está localizado entre dos genes que se transcriben en direcciones opuestas, *C8orf13* y *BLK*. Este polimorfismo se había descrito en el GWAS de Hom y col., con un efecto similar al encontrado (OR = 1.39) (Hom G y col., 2008). Otros dos estudios encontraron asociaciones significativas en este *locus*, uno de ellos con un SNP en *BLK* fuertemente ligado a rs13277113 (Graham RR y col., 2008), y el otro con un SNP independiente (SLEGEN y col., 2008). Estos resultados sugieren la posibilidad de la existencia de dos efectos genéticos independientes en la región. El *locus* también se ha visto asociado recientemente con AR (Gregersen PK y col., 2009) y esclerosis sistémica (Gourh P y col., 2010). No se ha identificado ningún polimorfismo que tuviese un efecto funcional. Aún así, se observó que el alelo de riesgo de rs13277113 se correlaciona con niveles bajos de mRNA de *BLK* y altos de *C8orf13*. Por lo tanto, cualquiera de estos dos efectos podría tener un papel en el LES. Sin embargo, mientras que *C8orf13* es un gen de función desconocida, *BLK* es un buen candidato en el LES porque está implicado en la ruta de señalización del BCR y en la diferenciación de células pro-B a pre-B (Tretter T y col., 2003). Además, como se comentó anteriormente, existen evidencias de que BLK interacciona con otro factor de susceptibilidad al

LES, BANK1, que interviene igualmente en la ruta de señalización del BCR (Castillejo-Lopez C y col., 2009). Esta interacción se observó tanto desde el punto de vista estadístico como físico, detectando la co-localización y co-inmunoprecipitación entre estas dos moléculas y de ellas con el receptor IP₃R, importante en las rutas de señalización. Por último, se ha observado que el haplotipo de riesgo de rs13277113 y de otro SNP de la región estaría relacionado con un defecto en la señalización del BCR y de la ruta NFκB (Gourh P y col., 2010). Por lo tanto, parece muy posible que el efecto de esta asociación esté relacionado con *BLK* y no con *C8orf13*.

1.2. Análisis funcionales del SNP asociado a LES, PD1.3

Uno de los factores genéticos de mayor interés en el LES hasta hace poco fue *PDCDI*, dado su papel crítico como inmunoregulator en múltiples enfermedades y su localización en un *locus* de ligamiento (Prokunina L y col., 2002; Keir ME y col., 2007; Day CL y col., 2006; Urbani S y col., 2006). El SNP PD1.3, localizado en el cuarto intrón de este gen, fue propuesto como el polimorfismo causal responsable de la asociación con LES (Prokunina L y col., 2002). Esto se debe a que se encuentra en una región con características compatibles con un *enhancer*. La región presenta múltiples sitios de unión a factores de transcripción relacionados con funciones inmunes y 4 repeticiones en *tandem* de unos 40 nucleótidos. La única evidencia experimental de que PD1.3 puede repercutir funcionalmente sobre el gen *PDCDI* era un EMSA. En este ensayo se veía alterada la unión del factor de transcripción RUNX1, relacionado con la diferenciación hematopoyética (Prokunina L y col., 2002).

Para estudiar el hipotético papel funcional del SNP PD1.3 y el posible papel de la región como *enhancer* se realizaron estudios de gen *reporter* en los

que se analizaron los efectos transcripcionales de construcciones plasmídicas que incluían el intrón 4 de *PDCD1*. Se utilizaron 4 construcciones para cada uno de los alelos, con el intrón 4 en las dos orientaciones, tanto en posición 5' como 3' del gen *reporter*, ya que los *enhancer* son independientes de su posición y orientación. Los resultados no mostraron evidencias de que el intrón 4 funcione como *enhancer*. Además, tampoco se observó un efecto diferencial de los dos alelos de PD1.3.

Estos resultados negativos pusieron en duda la importancia de la unión de RUNX1 a PD1.3. Así pues, se repitió el EMSA para comprobar la unión entre RUNX1 y oligonucleótidos sintéticos con las dos variantes alélicas de PD1.3. Los resultados confirmaron la unión preferencial del alelo G. Sin embargo, al comparar esta unión con la de las secuencias control, se pudo ver que es una unión ineficiente, a pesar de presentar la misma secuencia consenso. Esto llevó a sospechar que los nucleótidos vecinos a la secuencia consenso podían tener un papel en la afinidad de RUNX1 por el oligonucleótido. Por ello, se observaron las secuencias sintéticas que se habían unido a RUNX1 en un ensayo de afinidad y a partir de las cuales Meyers y col. describieron el sitio consenso de este factor de transcripción (TGT/cGGT) (Meyers S y col., 1993). Se comprobó que las bases adenina y timina son más frecuentes que las bases guanina y citosina, que son las que se encuentran en torno a PD1.3. Esta hipótesis fue confirmada mediante EMSA, al modificar las sondas originales en dichas posiciones. Otros autores ya habían evidenciado que los nucleótidos que limitan la secuencia consenso podían intervenir en la unión de RUNX1 al ADN, lo que apoyaba nuestros resultados. Así, sugirieron que la secuencia que mejor unía RUNX1 era el decanucleótido TTTGCGGTTA/T (Thornell A y col., 1991). Ésta es la misma secuencia que está presente en dos de los oligonucleótidos con

los que se observó una mejor afinidad de unión en nuestros experimentos: el control C₁ y el oligonucleótido de *PDCDI* modificado, P₁₂^m.

Estos resultados muestran las limitaciones de la identificación bioinformática de los lugares de unión a factores de transcripción. Ésta se basa en secuencias consenso (Roulet E y col., 1998) o bien en matrices de peso posicional que recogen la frecuencia de cada nucleótido en cada posición (Stormo GD y col., 2000). Sin embargo, carecen de información cuantitativa de la afinidad de unión del factor de transcripción por dicha secuencia. Además, la especificidad de los sitios de unión puede ser pequeña, ya que la mayoría de factores de transcripción son miembros de familias cuyos integrantes reconocen secuencias similares (Rahmann S y col., 2003). En consecuencia, *softwares* como TRANSFAC® Professional/Match™ (Biobase Biological Databases) (Kel AE y col., 2003) predicen un número elevado de posibles lugares de unión, de los que son funcionales sólo unos pocos (Li Q y col., 2001). De hecho, al hacer una búsqueda en TRANSFAC® Professional/Match™ de lugares de unión para RUNX1, se encontraron 70 secuencias que podrían unir este factor de transcripción a lo largo del gen *PDCDI*, de las cuales 11 estaban en el intrón 4.

Otra razón para la falta de especificidad de los lugares de unión predichos es que, en muchos casos, las secuencias de unión a factores de transcripción son identificadas a partir de ensayos *in vitro*, como EMSAs. Estos ensayos, al no estar dentro del contexto celular, tienen una utilidad limitada. Esto se debe a que no tienen en cuenta otra serie de variables que pueden determinar la unión del factor de transcripción y, por tanto, la funcionalidad de la región. Tales variables pueden ser interacciones proteína-proteína o modificaciones de la cromatina independientes de la secuencia (como metilaciones, posición de los

nucleosomas o modificaciones de histonas). Por todo ello, la identificación de las dianas funcionales de los factores de transcripción es todavía un reto sin resolver (Li CC y col., 2006). Un avance en la identificación de estas dianas consenso, que permite su evaluación *in vivo*, se ha obtenido con los estudios de genoma completo que combinan inmunoprecipitación de cromatina (ChIP) con la hibridación en *microarrays* (ChIP-on-chip) o con secuenciación (Hollenhorst PC y col., 2007; Yochum GS y col., 2007). Estos estudios detectan lugares de unión más grandes que los hexanucleótidos que se solían describir con las técnicas *in vitro*. Además, tienen en cuenta la influencia de nucleótidos vecinos, lo que permite una mayor especificidad de las secuencias identificadas. Desafortunadamente, los estudios de este tipo realizados hasta el momento no incluyen RUNX1.

Por otra lado, PD1.3 se encuentra en una región con repeticiones en *tandem*, característica que se ha observado en algunos *enhancer* y que podría apoyar un posible papel funcional para este polimorfismo (MacKenzie A y col., 1999; Giambra V y col., 2005; Bellizi D y col., 2005). Sin embargo, los polimorfismos que alteran la regulación transcripcional llevada a cabo por estas secuencias son frecuentemente variaciones en el número de copias de dichas repeticiones en *tandem*. Es menos probable que un SNP que altere uno de los sitios de unión de un factor de transcripción en una estructura repetitiva tenga consecuencias funcionales, ya que sería compensado por la unión de los mismos factores de transcripción a las otras repeticiones. Desde otro punto de vista, si la secuencia donde se encuentra PD1.3 fuese un elemento de regulación de *PDCDI*, habría sido sometida a selección positiva por su relevancia biológica. Por tanto, sería una secuencia conservada a lo largo de la evolución de las especies (Miller W y col., 2007). Éste no es el caso del intrón 4 de *PDCDI*, donde no hemos visto ninguna homología con genomas de mamíferos. Aún así,

esto no excluiría una posible relevancia funcional, ya que una fracción de los elementos funcionales no se han conservado. Otra característica de los *enhancer* típicos es que se encuentran en posición 5' ó 3' al gen o en el primer intrón del mismo, no habiéndose descrito en otros intrones. Así pues, estas observaciones junto con los resultados obtenidos en los ensayos de gen *reporter* y en los EMSAs cuestionan la relevancia funcional de este sitio en la regulación de la expresión de *PDCDI* y en la susceptibilidad al LES. Al excluirse PD1.3, queda abierta la posibilidad de que otro polimorfismo de *PDCDI* fuese el responsable de la susceptibilidad. Sin embargo, no hay ningún candidato claro y los GWAS de LES publicados recientemente no mostraron ninguna señal de asociación en este *locus*, por lo que en la actualidad no está confirmado que *PDCDI* sea un factor de susceptibilidad al LES.

1.3. Influencia del componente genético en la variabilidad clínica del LES

El LES es una enfermedad que muestra una gran heterogeneidad clínica por razones desconocidas. Sin embargo, parece que el componente genético contribuye a la variedad de fenotipos (Wallace DJ y col., 2007; Tsao BP y col., 2004; Sestak AI y col., 2007; Chung SA y col., 2009; Tsao BP y col., 2002; Scofield RH y col., 2003; Uribe AG y col., 2004; Ramos PS y col., 2006; Peschken CA y col., 2009). Este efecto no se puede demostrar en los estudios de asociación caso-control, en los que se tiene en cuenta al conjunto de los pacientes. De ahí el interés del análisis genético estratificando por las variables clínicas. Por otro lado, se han identificado dos subpoblaciones europeas bien diferenciadas desde el punto de vista de los polimorfismos genéticos, correspondientes a los países del norte y a los del sur (Seldin MF y col., 2006;

Novembre J y col., 2008; Tian C y col., 2008). Esta distinción entre subpoblaciones europeas puede influenciar a la variabilidad fenotípica del LES.

En primer lugar, los 9 *loci* asociados con LES que se han estudiado (*ITGAM*, *KIAA1542*, *PXK*, *1q25.1*, *C8orf13-BLK*, *STAT4*, *TYK2*, *MECP2* y *BANK1*) se analizaron estratificando por las manifestaciones clínicas (10 fenotipos definidos por los criterios de clasificación de la ACR y edad de aparición de la enfermedad). Las asociaciones observadas fueron débiles y poco consistentes. La asociación más clara fue la del alelo menor del SNP rs13277113 de *C8orf13-BLK* con mayor riesgo de nefritis. Por otra parte, *STAT4* es el factor genético que ha mostrado asociaciones más sólidas con variables clínicas en varios estudios. Sin embargo, el alelo menor del SNP rs7574865 de *STAT4* sólo mostró asociación con edad temprana de la enfermedad y una tendencia de asociación con menor susceptibilidad a úlceras orales. La escasez de asociaciones consistentes se debe, en parte, a las limitaciones de este tipo de análisis.

Una de las limitaciones de estos análisis de subgrupos consiste en la reducción del tamaño muestral, que disminuye la probabilidad de encontrar asociaciones significativas (Guillemin F, 2007; Brookes ST y col., 2001; Moyé LA, 2003). Otra limitación es el elevado número de comparaciones, que favorece la aparición de asociaciones debidas al azar. Además, la recogida retrospectiva de los datos y la variación existente entre las características clínicas de distintas colecciones de muestras (Taylor KE y col., 2008; Hom G y col., 2008; Cervera R y col., 2006; Pons-Estel BA y col., 2004; Alarcón GS y col., 2002; Petri M y col., 1997; Wang F y col., 1997; Al Arfaj AS y col., 2009) pueden generar asociaciones falsas (Hulley SB, 2007). A pesar de todos estos inconvenientes hay motivos para el optimismo, pues hay publicaciones con

resultados prometedores que asocian factores de susceptibilidad al LES con algún aspecto específico de su clínica (Taylor KE y col., 2008; Sigurdsson S y col., 2008^b; Tsao BP y col., 2004; Lauwerys BR y col., 2005; Karassa FB y col., 2003; Sestak AL y col., 2007).

Los problemas derivados del análisis de subgrupos se trataron de evitar siguiendo las recomendaciones realizadas en el campo de los ensayos clínicos (Guillemin F, 2007; Brookes ST y col., 2001; Moyé LA, 2003). Estas recomendaciones intentan obtener un equilibrio adecuado entre la tasa de falsos positivos y falsos negativos. Para obtenerlo, se utiliza una aproximación secuencial; se considera la consistencia y el poder estadístico de los análisis para la interpretación de los resultados; y se realizan correcciones por el número de comparaciones. La aproximación secuencial supone que sólo se estudia un subgrupo de factores de un grupo que ya mostró asociación. Por ello, se decidió evaluar los 9 SNPs que están asociados con LES en nuestras muestras y que no habían sido analizadas en relación con la clínica en nuestros estudios previos (*ITGAM*, *KIAA1542*, *PXK*, *1q25.1*, *C8orf13-BLK*, *STAT4*, *TYK2*, *MECP2* y *BANK1*). Esta es una aproximación práctica, pero no se puede excluir que un factor genético pueda estar asociado con un fenotipo y no con la enfermedad en su conjunto. Esto ocurre, por ejemplo, con el polimorfismo no sinónimo *FCGR3A* V158F y la nefritis lúpica (Karassa FB y col., 2003). Sin embargo, estas asociaciones podrán descubrirse más fácilmente en el contexto de GWAS o estudios de genes candidato en los que se analice de forma específica un fenotipo particular. Por otra parte, la consistencia de las asociaciones se evaluó mediante comprobación de su presencia en nuevos subgrupos de pacientes realizados de forma independiente. Estos subgrupos se pueden realizar de forma aleatoria, pero se prefirió utilizar una variable como el sexo, que puede estar relacionada con sesgos y que es un criterio comúnmente utilizado para dividir

pacientes con LES. Los resultados también se interpretaron teniendo en cuenta la potencia estadística de cada comparación SNP-fenotipo por separado. Esta potencia depende de la frecuencia alélica del SNP y de la frecuencia de la variable clínica. Los 9 SNPs tienen frecuencias alélicas similares (15-35 %), por lo que su efecto en la potencia estadística fue similar. Sin embargo, los criterios clínicos determinan una gradación de la potencia estadística. Ésta es mayor para los criterios con frecuencia entorno al 50 %, ya que permiten dividir a los casos en dos grupos de tamaño similar. Entre ellos se encuentran la edad de aparición de la enfermedad, la fotosensibilidad o el eritema malar. Sin embargo, los criterios con frecuencias alélicas más extremas son los que determinan potencias estadísticas menores. Entre estos criterios, se encuentran el desorden neurológico (de baja frecuencia), la artritis y el sexo femenino (ambas de alta frecuencia).

La asociación más clara y consistente, como ya se comentó, fue la del SNP rs13277113 en *C8orf13-BLK* con nefritis lúpica, ya que persistió tras corregir por el número de comparaciones y en los subgrupos de hombres y mujeres. El mismo SNP también se había evaluado en relación con la clínica en un GWAS. En dicho estudio, no se encontró ninguna asociación significativa, y además el efecto tendía a ser opuesto al encontrado aquí en relación con la nefritis (Hom G y col., 2008). Este resultado discordante puede estar relacionado con la notable diferencia en la prevalencia de la nefritis entre los dos estudios. En el GWAS, la prevalencia era significativamente menor (28 %) que en nuestra colección de muestras (41 %) ($P = 2 \times 10^{-12}$). Por lo tanto, la asociación de este SNP en *C8orf13-BLK* con nefritis necesita de estudios adicionales antes de considerarla como firme.

Una segunda asociación, aunque menos clara, fue la del SNP rs17435 de *IRAK1-MECP2* con nefritis. Se observó tanto en el análisis global, como de forma consistente en mujeres y hombres. Sin embargo, esta asociación no persistió tras realizar la corrección por múltiples tests. Por lo tanto, es necesario confirmarla en estudios futuros.

El rs7574865 de *STAT4* es el SNP que ha mostrado asociación más clara con la clínica, en concreto, con un fenotipo de LES severo. En un estudio se observó la asociación de *STAT4* con mayor riesgo de nefritis, aparición de la enfermedad en menores de 30 años, presencia de desorden inmunológico, en concreto, anti-ADN de doble cadena, y con ausencia de úlceras orales (Taylor KE y col., 2008). Otro estudio confirmó la asociación con presencia de daño inmunológico y anti-ADN de doble cadena y, de forma modesta, con riesgo de nefritis (Sigurdsson S y col., 2008^b). Nuestros resultados no confirmaron de forma clara ninguno de los datos anteriores. Sólo mostraron una débil asociación con aparición temprana de la enfermedad, que se mantenía en los subconjuntos de hombres y mujeres. También se observó una tendencia de asociación con ausencia de úlceras orales, que sólo era significativa en mujeres. Los análisis de *STAT4* con nefritis y daño inmunológico tenían una buena potencia estadística, pero no confirmaron los resultados previos. La combinación de estos datos con los de Taylor y col. reforzó la asociación de *STAT4* con daño inmunológico y con la ausencia de úlceras orales. Por consiguiente, este factor genético parece estar asociado con un fenotipo de LES particular. Éste podría incluir la presencia de daño inmunológico y, específicamente, anti-ADN de doble cadena, con los que está asociado en los dos estudios previos (Taylor KE y col., 2008; Sigurdsson S y col., 2008^b). También incluiría la aparición temprana de la enfermedad, con la cual se

observan asociaciones débiles en nuestro estudio y el de Taylor y col. y una tendencia no significativa en un tercero (Sigurdsson S y col., 2008^b).

Otro de los SNPs más fuertemente asociados con LES, el nsSNP rs1143679 de *ITGAM*, mostró una débil asociación con la edad temprana de aparición de la enfermedad. Sin embargo, al corregir por el número de comparaciones la asociación deja de ser significativa y, estratificando por sexo, sólo se encuentra asociado en mujeres. Además, esta asociación no se observó en estudios previos, por lo que es de dudosa relevancia. Al combinar nuestros datos con los de las variables que están asociadas con *ITGAM* en el estudio de Kim-Howard y col. aparece una diferencia débilmente significativa de *ITGAM* con daño inmunológico y, menos concluyente, con nefritis y eritema discoide. Un tercer trabajo que analizó *ITGAM* en relación con los criterios clínicos no encontró ninguna asociación significativa (Hom G y col., 2008). En este estudio se analizó un SNP que está sólo parcialmente relacionado con el rs1143679, por lo que no se pudieron utilizar sus datos para el análisis combinado.

Del resto de *loci* estudiados, sólo *TYK2* y *BANK1* fueron previamente analizados en relación con la clínica, sin ningún resultado de asociación significativa (Sigurdsson S y col., 2005; Guo L y col., 2009). Nuestros resultados son concordantes ya que no se encontró ninguna asociación para estos dos factores genéticos, a pesar de tener una buena potencia estadística.

En conclusión, se muestra una escasez de resultados sólidos de asociación de factores de susceptibilidad con fenotipos específicos. Esto se observa con los *loci* estudiados (Taylor KE y col., 2008; Sigurdsson S y col., 2008^b; Hom G y col., 2009; Guo L y col., 2009; Sigurdsson S y col., 2004) y también con otros *loci* confirmados en LES, como *IRF5* (Sigurdsson S y col., 2005; Cunninghame

Graham DS y col., 2007^b; Sigurdsson S y col., 2008^a; Ferreiro-Neira I y col., 2007), *PTPN22* (Reddy MV y col., 2005; Kyogoku C y col., 2004; Orozco G y col., 2005; Balada E y col., 2006) o *TNFAIP3* (Bates JS y col., 2009). El *HLA* únicamente se ha relacionado con la producción de anticuerpos específicos (Galeazzi M y col., 2002; Galeazzi M y col., 2000; Vasconcelos C y col., 2009; Fernando MM y col., 2007). Por tanto, parece no haber una correlación consistente entre los alelos de susceptibilidad en su conjunto y las variables clínicas del LES. El factor que muestra una evidencia más clara de asociación con la clínica es *STAT4*, aunque aún hay discrepancias que no permiten definir con exactitud el subgrupo de pacientes con LES asociado con polimorfismos en este gen. De ahí la necesidad de llevar a cabo estudios más grandes que permitan analizar un mayor número de pacientes con y sin cada una de las variables clínicas. Estos estudios aumentarán la probabilidad de encontrar asociaciones modestas o débiles que no se han detectado en los estudios actuales. Además, permitirán identificar subgrupos de pacientes más uniformes, lo que podría resultar útil a la hora de mejorar el tratamiento de la enfermedad.

Otra aproximación para explicar la variabilidad clínica del LES consiste en evaluar la influencia del componente genético de las subpoblaciones dentro del grupo étnico estudiado. Un trabajo previo mostró, con esta aproximación original, que parte de la heterogeneidad clínica en caucásicos depende de factores genéticos relacionados con la ascendencia nor- o sureuropea (Chung SA y col., 2009). Nuestros resultados confirmaron la mayor prevalencia de fotosensibilidad y menor prevalencia de artritis en el norte de Europa. Además, proporcionaron información complementaria a la aportada por Chung y col. debido al diferente diseño de los dos estudios. En el estudio de Chung y col. las muestras, de origen europeo, fueron recogidas en Estados Unidos, mientras que nuestras muestras han sido recogidas en Europa. En el estudio de Chung y col.

se utilizaron marcadores genéticos informativos de ascendencia (Seldin MF y col., 2006; Tian C y col., 2008). Mientras que en nuestro estudio se utilizó el país de origen como indicador de ascendencia, según estudios previos que observaron que la estructura genética europea refleja la distribución geográfica de los países de origen (Novembre J y col., 2008). Además, los marcadores de ascendencia europea se identificaron a partir de muestras de distintos países recogidos de un modo similar al de nuestro estudio (Seldin MF y col., 2006; Tian C y col., 2008). Otra discrepancia es que en el estudio de Chung y col. predominaban las muestras con ascendencia noreuropea, mientras que nuestro estudio presenta un número relativamente mayor de muestras del sur.

La diferencia, señalada en el párrafo anterior, en cuanto al lugar de recogida de las muestras es importante porque cada colección de pacientes puede estar afectada por factores locales, de tipo ambiental y socioeconómico, muy diferentes en Europa y Estados Unidos. La distinta recogida y recuperación de los datos clínicos o el tipo de pacientes que acuden a los centros sanitarios también podrían tener un efecto en la gran variabilidad clínica entre las distintas colecciones o estudios (Alarcón GS y col., 2002; Cervera R y col., 2006; Petri M, 1997; Wang F y col., 1997; Hom G y col., 2008; Taylor KE y col., 2008). Esta variabilidad disminuye la reproducibilidad de los estudios de asociación con factores genéticos. Para intentar evitar este problema sólo se analizaron aquellas características clínicas con menor variabilidad entre las colecciones de muestras del norte y entre las colecciones de muestras del sur. Específicamente, con un coeficiente de variación (CV) menor del 30 %. Éste es el menor CV observado para las variables clínicas de LES entre grupos étnicos diferentes (Alarcón GS y col., 2002; Cervera R y col., 2006; Petri M, 1997; Wang F y col., 1997), por lo que dentro de un mismo grupo étnico se debería esperar una menor variabilidad, aunque no mucho menor. A pesar de las diferencias entre

los estudios, la replicación de dos de las asociaciones observadas por Chung y col., como se comentó anteriormente, da fiabilidad a los resultados.

El interés de estos resultados reside en poder encontrar los mecanismos subyacentes que expliquen la heterogeneidad clínica del LES. Así, Chung y col. dieron como posible explicación de la mayor prevalencia de fotosensibilidad en el norte de Europa, el color de piel más claro, característico de estas poblaciones y que viene determinado genéticamente (Chung SA y col., 2009; Sestak AL y col., 2008). Además, esta explicación estaría en concordancia con una menor prevalencia de fotosensibilidad en los pacientes con LES afro-americanos (Sestak AL y col., 2008).

Aparte de analizar la influencia del componente genético en la variabilidad de los criterios de clasificación de la ACR, sería interesante realizar un estudio similar con otros datos clínicos de la enfermedad, como los índices de daño o actividad del LES. El problema para realizar estos análisis está en la dificultad que supone disponer de colecciones de muestras bien caracterizadas para todas estas variables. Este tipo de estudios requieren de un gran esfuerzo por parte de los centros que coleccionan las muestras, como se ha intentado hacer en el Proyecto Euro-Lupus, que ha recogido datos clínicos de una gran cantidad de pacientes europeos de forma prospectiva (Cervera R y col., 2006).

1.4. Componente genético común a LES y AR

En las últimas décadas, se han identificado una serie de genes que confieren susceptibilidad a varias enfermedades autoinmunes. La asociación del *HLA* es la más consistentemente establecida en estas enfermedades. *PTPN22*, confirmado como un factor de riesgo para AR (Begovich AB y col., 2004),

también confiere susceptibilidad a otras enfermedades autoinmunes (Gregersen PK y col., 2006^b) como T1D (Bottini N y col., 2004; Onengut-Gumuscu S y col., 2004), enfermedad tiroidea autoinmune (Velaga MR y col., 2004), algunas formas de artritis juvenil (Hinks A y col., 2005, Viken MK y col., 2005) y LES (Kyogoku C y col., 2004; Reddy MV y col., 2005; Wu H y col., 2005; Kaufman KM y col., 2006). *STAT4* también se observó asociado con AR y LES de forma reproducible (Remmers EF y col., 2007), mientras que su asociación con SS (Korman BD y col., 2008), T1D (Martinez A y col., 2008; Zervou MI y col., 2008^a) y enfermedad inflamatoria intestinal (Martinez A y col., 2008) necesita ser confirmada. La asociación de *IRF5* con AR y LES también ha sido corroborada en los últimos años (Sigurdsson S y col., 2005; Dieguez-Gonzalez R y col., 2008). Más recientemente se ha observado que la región *6q23* sería otro *locus* de susceptibilidad implicado en estas dos complejas enfermedades autoinmunes y en psoriasis (Wellcome Trust Case Control Consortium, 2007; Plenge RM y col., 2007^a; Plenge RM y col., 2007^b; Thomson W y col., 2007; Dieguez-Gonzalez R y col., 2009; Orozco G y col., 2009^b; Musone SL y col., 2008; Graham RR y col., 2008; Nair RP y col., 2009). Todas estas observaciones sugieren la existencia de un componente genético común a múltiples enfermedades autoinmunes y, por tanto, que dichas enfermedades compartan rutas patogénicas (Gregersen PK y col., 2006^a; Becker KG y col., 2004; Becker KG y col., 1998; Griffiths MM y col., 1999; Raman K y col., 2003). De ahí que los factores genéticos que se encuentran asociados con una enfermedad sean considerados como candidatos para otras enfermedades del mismo tipo. El inconveniente de esta estrategia es que se puede introducir un sesgo, ya que el estudio del mismo factor genético en múltiples enfermedades puede dar lugar a alguna asociación por azar.

De los 9 *loci* asociados con LES que se han estudiado (*STAT4*, *ITGAM*, *C8orf13-BLK*, *TYK2*, *1q.25.1*, *PXK*, *KIAA1542*, *MECP2* y *BANK1*), sólo *STAT4* había sido ampliamente analizado y claramente asociado con AR (Wellcome Trust Case Control Consortium, 2007; Lee HS y col., 2007; Remmers EF y col., 2007; Barton A y col., 2008; Kobayashi S y col., 2008; Orozco G y col., 2008; Martinez A y col., 2008; Zervou MI y col., 2008^b; Palomino-Morales RJ y col., 2008). También se confirmó este resultado en nuestra colección de muestras de AR. Sin embargo, no se encontró ninguna asociación significativa para los otros 8 factores genéticos, a pesar de que el poder estadístico de este estudio era suficiente para detectar efectos con OR de 1.15-1.20 de SNPs en cromosomas autosómicos y 1.24 del SNP de *MECP2*, en el cromosoma X. Además, estos efectos serían similares a las OR más débiles encontradas en las dos grandes colecciones de muestras de AR, NARAC (*North American Rheumatoid Arthritis Consortium*) y EIRA (*Swedish Epidemiological Investigation of Rheumatoid Arthritis*) (Plenge RM y col., 2005).

La disponibilidad de otros estudios de asociación con AR de los SNPs analizados, excepto para los de *ITGAM* y *MECP2*, permitió realizar un meta-análisis con mayor poder estadístico. En el mismo, se confirmó la asociación de *STAT4*. Además, se encontraron dos asociaciones de pequeño efecto en los *loci* *C8orf13-BLK* (OR = 1.09) y *PXK* (OR = 1.06), que no habían sido detectadas en los estudios previos. Recientemente, otro grupo también ha observado asociación de *C8orf13-BLK* con AR (Gregersen PK y col., 2009). Por tanto, los SNPs de este grupo asociados con AR parecen tener un efecto mucho menor que en LES, en donde mostraban ORs superiores a 1.19 (salvo *TYK2* y *LY9*, que no mostraron asociación en algún estudio previo (Cunningham Graham DS y col., 2007^a) o en nuestro estudio caso-control de LES). Incluso el efecto de *STAT4* es menor en AR que el que tiene en LES (OR = 1.14-1.32 vs 1.55-1.62).

Así pues, parece poco probable que estos SNPs tengan una importancia similar en ambas enfermedades.

En parte, los resultados de los 9 *loci* de susceptibilidad a LES eran de esperar, ya que la mayoría fueron encontradas en los GWAS de LES (SLEGEN y col., 2008; Kozyrev SV y col., 2008; Hom G y col., 2008; Graham RR y col., 2008), pero no en los de AR (Raychaudhuri S y col., 2008; Plenge RM y col., 2007^b; Wellcome Trust Case Control Consortium, 2007). Otra posible causa de la falta de una asociación clara de estos SNPs con AR es que distintos polimorfismos del mismo gen estén implicados en enfermedades distintas, como podría ocurrir con *BANK1*. Un estudio reciente de este gen ha encontrado asociación con AR, pero no con el SNP de LES estudiado aquí (rs17266594) (Orozco G y col., 2009^a).

Algo similar parece ocurrir con *TNFAIP3*. Por un lado, los SNPs de baja frecuencia, rs5029939, rs2230926 y rs7749323, fuertemente asociados con LES ($OR \geq 2$, $P \leq 10^{-4}$) (Musone SL y col., 2008; Graham RR y col., 2008), no mostraron asociación con AR. Por otro lado, el SNP rs610604 presentó una asociación con AR cuyo efecto era opuesto al observado en psoriasis (Nair RP y col., 2009). Sin embargo, dicho SNP junto con el rs644340 mostraban un efecto protector similar al del SNP que ya se había estudiado con anterioridad (rs582757) (Dieguez-Gonzalez R y col., 2009) y con la que tienen cierta relación. Por tanto, se confirmó dicha señal protectora en el gen *TNFAIP3*. Como ya se había mostrado previamente, ésta contribuye a la susceptibilidad a AR de forma conjunta con las señales protectora (rs13207033, rs10499194) y de riesgo (rs6920220) localizadas en la región intergénica de *6q23*. Así pues, el patrón de asociación con AR en *TNFAIP3* parece distinto al de LES o psoriasis. Sin embargo, estos resultados en AR necesitan ser confirmados en estudios

independientes debido a que las señales de asociación en *TNFAIP3* son débiles y no concuerdan con los resultados de otros dos estudios recientes. Éstos muestran una asociación débil del nsSNP de LES rs2230926 tanto en población caucásica ($P = 0.03$, OR = 1.21) (Orozco G y col., 2009^b) como japonesa (Shimane K y col., 2010).

En cuanto a la otra señal de susceptibilidad al LES, localizada a ~200 kb 3' de *TNFAIP3*, en la región 5' del gen *PERP* (rs6922466) (Musone SL y col., 2008), los resultados obtenidos tampoco mostraron asociación con AR, a pesar de que el poder estadístico del estudio era suficiente para detectar los efectos observados en LES.

Una posibilidad de que en el análisis global de casos con AR vs controles no se detecte la asociación de la mayoría de factores de LES o de que haya discrepancias en las señales es que la asociación en AR tenga lugar en subgrupos de pacientes con características fenotípicas más parecidas al LES. Esto parece ocurrir en *IRF5*, cuya asociación con LES ha sido fuertemente confirmada (Sigurdsson S y col., 2005; Graham RR y col., 2007^a; Cunningham Graham DS y col., 2007^b; Kozyrev SV y col., 2007; Ferreiro-Neira I y col., 2007; Sigurdsson S y col., 2008^a), mientras que en AR está restringida a un subgrupo de pacientes que no está totalmente definido (Dieguez-Gonzalez R y col., 2008; Sigurdsson S y col., 2007). Sin embargo, el análisis de los 9 factores genéticos de LES estratificando por algunas variables clínicas de la AR sólo muestra una asociación consistente tras la corrección por tests múltiples. Esta asociación se corresponde con el *locus C8orf13-BLK* en relación con el SS. El alelo menor parece conferir protección frente a este fenotipo. Sin embargo, el mismo alelo también mostró mayor riesgo de AR en el análisis combinado con datos publicados, siendo este efecto opuesto de difícil explicación.

En relación con el *locus 6q23*, ya se había observado que la señal de protección en la región intergénica, marcada por el SNP rs13207033 (o su tagSNP rs10499194), estaba preferentemente asociada en los pacientes ACPA+ (Plenge RM y col., 2007^a; Dieguez-Gonzalez R y col., 2009). Sin embargo, en otro estudio se detectó una asociación con pacientes ACPA- (Orozco G y col., 2009^b). Al analizar variables clínicas más características del LES en estos pacientes con AR, se observó además que la asociación de este SNP es más marcada con pacientes ANA+. Sin embargo, el SNP rs610604 de *TNFAIP3* tenía una asociación más marcada con pacientes ANA-. Mientras que los SNPs de baja frecuencia característicos del LES seguían sin mostrar asociación tras esta estratificación. Estos resultados sugieren la posibilidad de que los pacientes con AR muestren un patrón de asociación particular en relación con la presencia de ANAs. Sin embargo, este patrón no parece relacionado con el observado en LES. Aún así, hay que tener en cuenta las limitaciones de estos análisis. Por una parte, no disponemos de los datos de ANAs para todos los pacientes con AR. Además, éstos fueron medidos mediante técnicas distintas y a tiempos diferentes en la evolución de la AR, pudiendo estar influenciados por el tratamiento con antagonistas del TNF, que se sabe que inducen la producción de estos autoanticuerpos (Charles PJ y col., 2000). En segundo lugar, el análisis de subfenotipos, como se comentó en el apartado anterior, implica una reducción del número de muestras, disminuyendo el poder estadístico del análisis. Por último, el alto número de comparaciones aumenta la probabilidad de aparición de falsos positivos. Una estrategia que permitiría aumentar la potencia estadística de los estudios genotipo-fenotipo, sería combinar datos similares de estudios independientes en un meta-análisis, pero de momento otros estudios no han presentado resultados de este tipo.

En conclusión, de los factores de susceptibilidad al LES estudiados, sólo *STAT4* es claramente compartido por la AR, mostrando la misma señal de riesgo asociada. *TNFAIP3* también sería un *locus* de susceptibilidad para estas dos enfermedades, aunque hemos observado que esta asociación no se debe a los mismos polimorfismos, por lo que posiblemente no estén implicados los mismos mecanismos. Además, los factores de susceptibilidad a LES *C8orf13-BLK* y *PXK* también han mostrado evidencias de estar asociados con AR. No obstante estos últimos resultados necesitan ser confirmados. Esta dificultad para obtener señales significativas en AR podría deberse en parte a que el efecto de los polimorfismos es, en general, más débil en AR que en LES. Aún así, otros estudios han mostrado más factores compartidos por LES y AR, como *PTPN22*, *IRF5*, *CTLA4* y *BANK1*. Por tanto, estas dos enfermedades comparten varios factores genéticos de susceptibilidad, aunque también existen factores específicos para cada una de ellas.

2. Sumario

En este trabajo se ha abordado la confirmación de factores genéticos de susceptibilidad al LES. Además, se ha realizado el análisis funcional de PD1.3, el posible SNP causal del gen *PDCDI*. También se ha evaluado la influencia del componente genético sobre los subfenotipos del LES. Y, por último, se ha explorado si los nuevos factores genéticos identificados en LES son compartidos por AR, en busca de mecanismos patogénicos comunes a estas dos enfermedades autoinmunes.

El estudio de nuevos factores de susceptibilidad al LES proporcionó la replicación independiente de la asociación de 9 *loci* con esta enfermedad. Como consecuencia de estos resultados, 5 de los *loci* fueron confirmados por primera

vez de modo independiente (*Iq25.1*, *MECP2*, *KIAA1542*, *PXK* y *BANK1*). Un sexto fue replicado, aunque su asociación sigue siendo controvertida (*TYK2*). De este modo, ya hay un número considerablemente alto de *loci* asociados de forma consistente con LES. Este alto grado de reproducibilidad es una mejora fundamental en la investigación de la genética del LES y otras enfermedades complejas y es consecuencia de los grandes estudios que se están llevando a cabo. Por lo tanto, estamos en una fase de descubrimientos interesantes en este campo. Sin embargo, la identificación de los factores genéticos es sólo un primer paso. Es necesario transformar la información que obtenemos en conocimiento útil y, de momento, tenemos pocas claves sobre el significado biológico de las asociaciones con LES detectadas. Estudios futuros deberán tratar de identificar las variantes causales y, mediante ensayos funcionales, determinar sus efectos a los niveles molecular, celular y de la enfermedad, como se ha intentado hacer con PD1.3.

Sin embargo, los resultados sobre PD1.3 cuestionan su supuesto papel funcional. Puede que este SNP sea relevante en algún sentido que no se haya explorado todavía, pero la información disponible no permite predecir ningún efecto. Además, su asociación con el LES también ha sido cuestionada, al menos en población europea, y en los GWAS no se ha observado ninguna señal significativa en *PDCDI*. Por lo tanto, se pone en duda el papel de este gen en la susceptibilidad al LES en caucásicos.

También se ha examinado el papel que los nuevos factores de susceptibilidad al LES tienen en los diferentes fenotipos de la enfermedad. Los resultados muestran una escasez de asociaciones destacables con características clínicas específicas. Las asociaciones más claras se observaron con *STAT4*, pero con un subfenotipo todavía no muy bien definido. Por tanto, la variabilidad

fenotípica no se correlaciona de modo consistente con los alelos de susceptibilidad al LES. Así, se hace necesario aumentar la potencia estadística de los estudios encaminados a descubrir polimorfismos con influencia en el fenotipo, para poder encontrar las posibles asociaciones débiles o modestas. Una posible aproximación sería el realizar GWAS o estudios de genes candidato en subgrupos de pacientes con un fenotipo concreto. También es posible que la falta de correlación con los alelos de susceptibilidad se deba a su intervención en interacciones más complejas, entre factores genéticos o de éstos con el ambiente. Una prueba adicional de que el componente genético influye en los fenotipos del LES se obtuvo al comparar pacientes con distinto fondo genético característico de las subpoblaciones de ascendencia nor- o sureuropea. Así, confirmamos una mayor prevalencia de fotosensibilidad y menor prevalencia de artritis en el norte que en el sur. Además, estas diferencias muestran la necesidad de tener en cuenta la ascendencia en los estudios epidemiológicos del LES.

Por último, observamos que una buena parte de los nuevos factores genéticos de LES eran específicos de LES, ya que no estaban asociados con AR. Así, el único factor de susceptibilidad de los 9 analizados que está claramente compartido por ambas enfermedades es *STAT4*. *C8orf13-BLK* y *PXK* mostraron una débil asociación con AR, pero sólo al combinar nuestros datos con los de otros estudios. En cuanto a *TNFAIP3*, nuestros resultados muestran señales de asociación distintas en las dos enfermedades. En resumen, los factores genéticos que son compartidos por el LES y la AR son *PTPN22*, *STAT4*, la región intergénica de *6q23* y puede que *IRF5*, *CTLA4*, *C8orf13-BLK*, *PXK* y *BANK1*. Sin embargo, el componente genético de ambas enfermedades difiere en muchos otros factores de susceptibilidad, como *TNFAIP3*, *KIAA1542*, *1q25.1*, *MECP2* y *TYK2*. Estos resultados indican que, aunque comparten una fracción

del componente genético, el resto del componente genético tiene un efecto específico de cada enfermedad. Si bien ésta no es una idea nueva, carecía de apoyo experimental. Sin embargo, todavía no se puede estimar la magnitud del componente diferencial ya que no se ha definido una gran parte de la heredabilidad de estas dos enfermedades. Esto será posible una vez que se hayan identificado las variantes causales de cada factor genético.

Concluyendo, nuestro trabajo ha contribuido al conocimiento del componente genético del LES, así como a diferenciarlo del de la AR. De esta forma, se ha contribuido al conocimiento de los mecanismos patogénicos que son importantes en estas enfermedades autoinmunes.

CONCLUSIONES

Conclusiones

1. Se han replicado de forma independiente nueve *loci* asociados a LES (*ITGAM*, *STAT4*, *C8orf13-BLK*, *1q25.1*, *MECP2*, *KIAA1542*, *PXK*, *BANK1* y *TYK2*). Seis de estas replicaciones tuvieron importancia confirmatoria, ya que fueron las primeras independientes (*1q25.1*, *MECP2*, *KIAA1542*, *PXK* y *BANK1*) o, en el caso de *TYK2*, porque su asociación era controvertida. Estos resultados muestran la reproducibilidad de los estudios genéticos en LES recientes, por lo que su futuro es prometedor.
2. La unión ineficiente a *RUNX1*, la ausencia de efecto *enhancer* y la falta de regulación transcripcional alelo-diferencial cuestionan el papel del SNP PD1.3 como polimorfismo causal de la asociación de *PDCD1* con LES.
3. Los efectos de los factores de susceptibilidad en la variabilidad clínica del LES son débiles y difíciles de reproducir. *STAT4* es el que está más claramente asociado, aunque el fenotipo todavía no está completamente definido.
4. Se confirma la mayor prevalencia de fotosensibilidad y la menor prevalencia de artritis en los pacientes con ascendencia del norte de Europa en relación con los de ascendencia del sur. Estos resultados corroboran la influencia del fondo genético en la heterogeneidad clínica del LES.

5. Ninguno de los *loci* asociados recientemente con LES, excepto *STAT4*, ha mostrado asociación clara con AR. Así mismo, las asociaciones con LES observadas en el gen *TNFAIP3* no coinciden con las observadas en AR. Por tanto, aunque AR y LES comparten una fracción de su componente genético, también existen factores de susceptibilidad específicos de cada enfermedad.

BIBLIOGRAFÍA

Bibliografía

- Abelson AK, Delgado-Vega AM, Kozyrev SV, Sanchez E, Velazquez-Cruz R, Eriksson N *et al.* *STAT4 associates with SLE through two independent effects that correlate with gene expression and act additively with IRF5 to increase risk.* Ann Rheum Dis. 2009; 68: 1746-1753.
- Aho K, Koskenvuo M, Tuominen J, Kaprio J. *Occurrence of rheumatoid arthritis in a nation wide series of twins.* J Rheumatol. 1986; 13: 899-902.
- Aiba Y, Yamazaki T, Okada T, Gotoh K, Sanjo H, Ogata M *et al.* *BANK negatively regulates Akt activation and subsequent B cell responses.* Immunity. 2006; 24: 259-268.
- Al Arfaj AS, Khalil N. *Clinical and immunological manifestations in 624 SLE patients in Saudi Arabia.* Lupus. 2009; 18: 465-473.
- Alarcón GS, Friedman AW, Straaton KV, Moulds JM, Lisse J, Bastian HM *et al.* *Systemic lupus erythematosus in three ethnic groups: III. A comparison of characteristics early in the natural history of the LUMINA cohort. LUPus in MInority populations: NAture vs. Nurture.* Lupus. 1999; 8: 197-209.
- Alarcon GS, McGwin G Jr, Petri M, Reveille JD, Ramsey-Goldman R, Kimberly RP. *Baseline characteristics of a multiethnic lupus cohort: PROFILE.* Lupus. 2002; 11: 95-101.
- Alarcón-Segovia D, Alarcón-Riquelme ME, Cardiel MH, Caeiro F, Massardo L, Villa AR *et al.* *Familiar aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort.* Arthritis Rheum. 2005; 52: 1138-1147.
- Arbuckle MR, McClain MT, Rubertone MV, Scofield RH, Dennis GJ, James JA *et al.* *Development of autoantibodies before the clinical onset of systemic lupus erythematosus.* N Engl J Med. 2003. 349: 1526-1533.

- Arita K, Ariyoshi M, Tochio H, Nakamura Y, Shirakawa M. *Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism*. Nature. 2008; 455: 818-821.
- Baechler EC, Gregersen PK, Behrens TW. *The emerging role of interferon in human systemic lupus erythematosus*. Curr Opin Immunol. 2004; 16: 801-807.
- Balada E, Villarreal-Tolchinsky J, Ordi-Ros J, Labrador M, Serrano-Acedo S, Martinez-Lostao L *et al*. *Multiplex family-based study in systemic lupus erythematosus: association between the R620W polymorphism of PTPN22 and the FcgammaRIIa (CD32A) R131 allele*. Tissue Antigens. 2006; 68: 432-438.
- Ballestar E, Esteller M, Richardson BC. *The epigenetic face of systemic lupus erythematosus*. J Immunol. 2006; 176: 7143-7147.
- Baranzini SE. *The genetics of autoimmune diseases: a networked perspective*. Curr Opin Immunol. 2009; 21: 596-605.
- Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA *et al*. *Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes*. Nat Genet. 2009; 41: 703-707.
- Barton A, Thomson W, Ke X, Eyre S, Hinks A, Bowes J *et al*. *Re-evaluation of putative rheumatoid arthritis susceptibility genes in the post-genome wide association study era and hypothesis of a key pathway underlying susceptibility*. Hum Mol Genet. 2008; 17: 2274-2279.
- Barton A, Worthington J. *Genetic Susceptibility to Rheumatoid Arthritis: An Emerging Picture*. Arthritis Rheum. 2009; 61: 1441-1446.
- Bates JS, Lessard CJ, Leon JM, Nguyen T, Battiest LJ, Rodgers J *et al*. *Meta-analysis and imputation identifies a 109 kb risk haplotype spanning TNFAIP3 associated with lupus nephritis and hematologic manifestations*. Genes Immun. 2009; 10: 470-477.
- Becker KG, Simon RM, Bailey-Wilson JE, Freidlin B, Biddison WE, McFarland HF *et al*. *Clustering of non-major histocompatibility complex*

susceptibility candidate loci in human autoimmune diseases. Proc Natl Acad Sci U S A. 1998; 95: 9979-9984.

- Becker KG. *The common variants/multiple disease hypothesis of common complex genetic disorders*. Med Hypotheses. 2004; 62: 309-317.
- Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC *et al*. *A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis*. Am J Hum Genet. 2004; 75: 330-337.
- Bellizzi D, Rose G, Cavalcante P, Covello G, Dato S, De Rango F *et al*. *A novel VNTR enhancer within the SIRT3 gene, a human homologue of SIR2, is associated with survival at oldest ages*. Genomics. 2005; 85: 258-263.
- Beyaert R, Heyninck K, Van Huffel S. *A20 and A20-binding proteins as cellular inhibitors of nuclear factor-kappa B-dependent gene expression and apoptosis*. Biochem Pharmacol. 2000; 60: 1143-1151.
- Blank C, Brown I, Marks R, Nishimura H, Honjo T, Gajewski TF *et al*. *Absence of programmed death receptor 1 alters thymic development and enhances generation of CD4/CD8 double-negative TCR-transgenic T cells*. J Immunol. 2003; 171: 4574-4581.
- Boh EE. *Neonatal lupus erythematosus*. Clin Dermatol. 2004; 22: 125-128.
- Boone DL, Turer EE, Lee EG, Ahmad RC, Wheeler MT, Tsui C *et al*. *The ubiquitin-modifying enzyme A20 is required for termination of Toll-like receptor responses*. Nat Immunol. 2004; 5: 1052-1060.
- Bottini N, Musumeci L, Alonso A, Rahmouni S, Nika K, Rostamkhani M *et al*. *A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes*. Nat Genet. 2004; 36: 337-338.
- Botto M, Walport MJ. *Clq, autoimmunity and apoptosis*. Immunobiology. 2002; 205: 395-406.

- Bristow CA, Shore P. *Transcriptional regulation of the human MIP-1alpha promoter by RUNX1 and MOZ*. Nucleic Acids Res. 2003; 31: 2735-2744.
- Brookes ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey Smith G. *Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives*. Health Technol Assess. 2001; 5: 1-56.
- Bruhn L, Munnerlyn A, Grosschedl R. *ALY, a context-dependent coactivator of LEF-1 and AML-1, is required for TCRalpha enhancer function*. Genes Dev. 1997; 11: 640-653.
- Cantor RM, Yuan J, Napier S, Kono N, Grossman JM, Hahn BH *et al*. *Systemic lupus erythematosus genome scan: support for linkage at 1q23, 2q33, 16q12-13, and 17q21-23 and novel evidence at 3p24, 10q23-24, 13q32, and 18q22-23*. Arthritis Rheum. 2004; 50: 3203-3210.
- Carreno BM, Collins M. *The B7 family of ligands and its receptors: new pathways for costimulation and inhibition of immune responses*. Annu Rev Immunol. 2002; 20: 29-53.
- Casciola-Rosen LA, Anhalt G, Rosen A. *Autoantigens targeted in systemic lupus erythematosus are clustered in two populations of surface structures on apoptotic keratinocytes*. J Exp Med. 1994; 179: 1317-30.
- Castaño-Rodríguez N, Diaz-Gallo LM, Pineda-Tamayo R, Rojas-Villarraga A, Anaya JM. *Meta-analysis of HLA-DRB1 and HLA-DQB1 polymorphisms in Latin American patients with systemic lupus erythematosus*. Autoimmun Rev. 2008; 7: 322-330.
- Castillejo-Lopez C, Delgado-Vega AM, Wojcik J, Kozyrev S, Egido JRL, Sanchez E *et al*. *Genetic Interactions Reveal a Novel B-Cell Pathway in Systemic Lupus Erythematosus [abstract]*. Arthritis Rheum. 2009; 60 Suppl 10: 1226.
- Cervera R, Abarca-Costalago M, Abramovicz D, Allegri F, Annunziata P, Aydintug AO *et al*. *Systemic lupus erythematosus in Europe at the change of*

the millennium: lessons from the "Euro-Lupus Project". Autoimmun Rev. 2006; 5: 180-186.

- Chahrour M, Jung SY, Shaw C, Zhou X, Wong ST, Qin J *et al.* *MeCP2, a key contributor to neurological disease, activates and represses transcription.* Science. 2008; 320: 1224-1229.
- Chang YK, Yang W, Zhao M, Mok CC, Chan TM, Wong RWS *et al.* *Association of BANK1 and TNFSF4 with systemic lupus erythematosus in Hong Kong Chinese.* Genes Immun. 2009; 10: 414-420.
- Charles PJ, Smeenk RJ, De Jong J, Feldmann M, Maini RN. *Assessment of antibodies to double-stranded DNA induced in rheumatoid arthritis patients following treatment with infliximab, a monoclonal antibody to tumor necrosis factor alpha: findings in open-label and randomized placebo-controlled trials.* Arthritis Rheum. 2000; 43: 2383-2390.
- Chowdhury D, Beresford PJ, Zhu P, Zhang D, Sung JS, Demple B *et al.* *The exonuclease TREX1 is in the SET complex and acts in concert with NM23-H1 to degrade DNA during granzyme A-mediated cell death.* Mol Cell. 2006; 23: 133-142.
- Chung SA, Criswell LA. *PTPN22: its role in SLE and autoimmunity.* Autoimmunity. 2007; 40: 582-590.
- Chung SA, Tian C, Taylor KE, Lee AT, Ortmann WA, Hom G *et al.* *European population substructure is associated with mucocutaneous manifestations and autoantibody production in systemic lupus erythematosus.* Arthritis Rheum. 2009; 60: 2448-2456.
- Cooper GS, Gilbert KM, Greidinger EL, James JA, Pfau JC, Reinlib L *et al.* *Recent advances and opportunities in research on Lupus: environmental influences and mechanisms of disease.* Environ Health Perspect. 2008; 116: 695-702.

- Coxon A, Rieu P, Barkalow FJ, Askari S, Sharpe AH, von Andrian UH *et al.* *A novel role for the beta 2 integrin CD11b/CD18 in neutrophil apoptosis: a homeostatic mechanism in inflammation.* Immunity. 1996; 5: 653-666.
- Coyle AJ, Gutierrez-Ramos JC. *The expanding B7 superfamily: increasing complexity in costimulatory signals regulating T cell function.* Nat Immunol. 2001; 2: 203-209.
- Crotty S, Johnston RJ, Schoenberger SP. *Effectors and memories: Bcl-6 and Blimp-1 in T and B lymphocyte differentiation.* Nat Immunol. 2010; 11: 114-120.
- Cunninghame Graham DS, Akil M, Vyse TJ. *Association of polymorphisms across the tyrosine kinase gene, TYK2 in UK SLE families.* Rheumatology (Oxford). 2007^a; 46: 927-930.
- Cunninghame Graham DS, Graham RR, Manku H, Wong AK, Whittaker JC, Gaffney PM *et al.* *Polymorphism at the TNF superfamily gene TNFSF4 confers susceptibility to systemic lupus erythematosus.* Nat Genet. 2008^a; 40: 83-89.
- Cunninghame Graham DS, Manku H, Wagner S, Reid J, Timms K, Gutin A *et al.* *Association of IRF5 in UK SLE families identifies a variant involved in polyadenylation.* Hum Mol Genet. 2007^b; 16: 579-591.
- Cunninghame Graham DS, Vyse TJ, Fortin PR, Montpetit A, Cai YC, Lim S *et al.* *Association of LY9 in UK and Canadian SLE families.* Genes Immun. 2008^b; 9:93-102.
- Day CL, Kaufmann DE, Kiepiela P, Brown JA, Moodley ES, Reddy S *et al.* *PD-1 expression on HIV-specific T cells is associated with T-cell exhaustion and disease progression.* Nature. 2006; 443: 350-354.
- Deapen D, Escalante A, Weinrib L, Horwitz D, Bachman B, Roy-Burman P *et al.* *A revised estimate of twin concordance in systemic lupus erythematosus.* Arthritis Rheum. 1992; 35: 311-318.

- Delgado-Vega AM, Abelson AK, Sánchez E, Witte T, D'Alfonso S, Galeazzi M *et al.* *Replication of the TNFSF4 (OX40L) promoter region association with systemic lupus erythematosus.* Genes Immun. 2009; 10: 248-253.
- D'Esposito M, Quaderi NA, Ciccodicola A, Bruni P, Esposito T, D'Urso M *et al.* *Isolation, physical mapping, and northern analysis of the X-linked human gene encoding methyl CpG-binding protein, MECP2.* Mamm Genome. 1996; 7: 533-535.
- Dieguez-Gonzalez R, Calaza M, Perez-Pampin E, Balsa A, Blanco FJ, Cañete JD *et al.* *Analysis of TNFAIP3, a feedback inhibitor of nuclear factor-kappa B and the neighbor intergenic 6q23 region in rheumatoid arthritis susceptibility.* Arthritis Res Ther. 2009; 11; R42.
- Dieguez-Gonzalez R, Calaza M, Perez-Pampin E, Rodriguez de la Serna A, Fernandez-Gutierrez B, Castañeda S *et al.* *Association of interferon regulatory factor 5 haplotypes, similar to that found in systemic lupus erythematosus, in a large subgroup of patients with rheumatoid arthritis.* Arthritis Rheum. 2008; 58: 1264-1274.
- Dieudé P, Wipff J, Guedj M, Ruiz B, Melchers I, Hachulla E *et al.* *BANK1 is a genetic risk factor for diffuse cutaneous systemic sclerosis and has additive effects with IRF5 and STAT4.* Arthritis Rheum. 2009; 60: 3447-3454.
- Diveu C, McGeachy MJ, Cua DJ. *Cytokines that regulate autoimmunity.* Curr Opin. Immunol. 2008; 20: 663-668.
- Dunne JL, Collins RG, Beaudet AL, Ballantyne CM, Ley K. *Mac-1, but not LFA-1, uses intercellular adhesion molecule-1 to mediate slow leukocyte rolling in TNF-alpha-induced inflammation.* J Immunol. 2003; 171: 6105-6111.
- Ehlers MR. *CR3: a general purpose adhesion-recognition receptor essential for innate immunity.* Microbes Infect. 2000; 2: 289-294.

- Fanciulli M, Norsworthy PJ, Petretto E, Dong R, Harper L, Kamesh L *et al.* *FCGR3B copy number variation is associated with susceptibility to systemic, but not organ-specific, autoimmunity.* Nat Genet. 2007; 39: 721-723.
- Fernando MM, Stevens CR, Sabeti PC, Walsh EC, McWhinnie AJ, Shah A *et al.* *Identification of two independent risk factors for lupus within the MHC in United Kingdom families.* PLoS Genet. 2007; 3: e192.
- Fernando MM, Stevens CR, Walsh EC, De Jager PL, Goyette P, Plenge RM *et al.* *Defining the role of the MHC in autoimmunity: a review and pooled analysis.* PLoS Genet. 2008; 4: e1000024.
- Ferreira-Neira I, Calaza M, Alonso-Perez E, Marchini M, Scorza R, Sebastiani GD *et al.* *Opposed independent effects and epistasis in the complex association of IRF5 to SLE.* Genes Immun. 2007; 8: 429-438.
- Ferreira-Vidal I, D'Alfonso S, Papasteriades C, Skopouli FN, Marchini M, Scorza R *et al.* *Bias in association studies of systemic lupus erythematosus susceptibility due to geographical variation in the frequency of a programmed cell death 1 polymorphism across Europe.* Genes Immun. 2007; 8: 138-146.
- Ferreira-Vidal I, Gomez-Reino JJ, Barros F, Carracedo A, Carreira P, Gonzalez-Escribano F *et al.* *Association of PDCD1 with susceptibility to systemic lupus erythematosus: evidence of population-specific effects.* Arthritis Rheum. 2004; 50: 2590-2597.
- Fielder AHL, Walport MJ, Batchelor JR, Rynes RI, Black CM, Dodi IA *et al.* *Family study of the major histocompatibility complex in patients with systemic lupus erythematosus: importance of null alleles of C4A and C4B in determining disease susceptibility.* Br Med J. 1983; 286: 425-428.
- Filaci G, Bacilieri S, Fravega M, Monetti M, Contini P, Ghio M *et al.* *Impairment of CD81 T suppressor cell function in patients with active systemic lupus erythematosus.* J Immunol. 2001; 166: 6452-6457.

- Finnegan A, Grusby MJ, Kaplan CD, O'Neill SK, Eibel H, Koreny T *et al.* *IL-4 and IL-12 regulate proteoglycan-induced arthritis through Stat-dependent mechanisms.* J Immunol. 2002; 169: 3345-3352.
- Forabosco P, Gorman JD, Cleveland C, Kelly JA, Fisher SA, Ortmann WA *et al.* *Meta-analysis of genome-wide linkage studies of systemic lupus erythematosus.* Genes Immun. 2006; 7: 609-614.
- Frank R, Zhang J, Uchida H, Meyers S, Hiebert SW, Nimer SD *et al.* *The AML1/ETO fusion protein blocks transactivation of the GM-CSF promoter by AML1B.* Oncogene. 1995; 11: 2667-2674.
- Franke A, Balschun T, Karlsen TH, Sventoraityte J, Nikolaus S, Mayr G *et al.* *Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility.* Nat Genet. 2008; 40: 1319-1323.
- Freeman GJ, Long AJ, Iwai Y, Bourque K, Chernova T, Nishimura H *et al.* *Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation.* J Exp Med. 2000; 192: 1027-1034.
- Fung EY, Smyth DJ, Howson JM, Cooper JD, Walker NM, Stevens H *et al.* *Analysis of 17 autoimmune disease-associated variants in type 1 diabetes identifies 6q23/TNFAIP3 as a susceptibility locus.* Genes Immun. 2009; 10: 188-191.
- Gaffney PM, Kearns GM, Shark KB, Ortmann WA, Selby SA, Malmgren ML *et al.* *A genome-wide search for susceptibility genes in human systemic lupus erythematosus sibpair families.* Proc Natl Acad Sci U S A. 1998; 95: 14875-14879.
- Gaffney PM, Ortmann WA, Selby SA, Shark KB, Ockenden TC, Rohlf KE *et al.* *Genome screening in human systemic lupus erythematosus: results from a second Minnesota cohort and combined analyses of 187 sib-pair families.* Am J Hum Genet. 2000; 66: 547-556.

- Galeazzi M, Sebastiani GD, Morozzi G, Carcassi C, Ferrara GB, Scorza R *et al.* *HLA class II DNA typing in a large series of European patients with systemic lupus erythematosus: correlations with clinical and autoantibody subsets.* Medicine (Baltimore). 2002; 81: 169-178.
- Galeazzi M, Sebastiani GD, Tincani A, Piette JC, Allegri F, Morozzi G *et al.* *HLA class II alleles associations of anticardiolipin and anti-beta2GPI antibodies in a large series of European patients with systemic lupus erythematosus.* Lupus. 2000; 9: 47-55.
- Gateva V, Sandling JK, Hom G, Taylor KE, Chung SA, Sun X *et al.* *A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus.* Nat Genet. 2009; 41: 1228-1233.
- Giambra V, Fruscalzo A, Giufre' M, Martinez-Labarga C, Favaro M, Rocchi M *et al.* *Evolution of human IgH30EC duplicated structures: both enhancers HSI,2 are polymorphic with variation of transcription factor's consensus sites.* Gene. 2005; 346: 105-114.
- Golden-Mason L, Klarquist J, Wahed AS, Rosen HR. *Cutting Edge: Programmed Death-1 Expression Is Increased on Immunocytes in Chronic Hepatitis C Virus and Predicts Failure of Response to Antiviral Therapy: Race-Dependent Differences.* J Immunol. 2008; 180: 3637-3641.
- Gourh P, Agarwal SK, Martin E, Divecha D, Rueda B, Bunting H *et al.* *Association of the C8orf13-BLK region with systemic sclerosis in North-American and European populations.* J Autoimmun. 2010; 34:155-62.
- Graham DS, Graham RR, Manku H, Wong AK, Whittaker JC, Gaffney PM *et al.* *Polymorphism at the TNF superfamily gene TNFSF4 confers susceptibility to systemic lupus erythematosus.* Nat Genet. 2008; 40: 83-89.
- Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM *et al.* *Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus.* Nat Genet. 2008; 40: 1059-1061.

- Graham RR, Hom G, Ortmann W, Behrens TW. *Review of recent genome-wide association scans in lupus*. J Intern Med. 2009; 265: 680-688.
- Graham RR, Kyogoku C, Sigurdsson S, Vlasova IA, Davies LR, Baechler EC *et al*. *Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus*. Proc Natl Acad Sci U S A. 2007^a; 104: 6758-6763.
- Graham RR, Ortmann W, Rodine P, Espe K, Largefeld C, Lange E *et al*. *Specific combinations of HLA-DR2 and DR3 class II haplotypes contribute graded risk for disease susceptibility and autoantibodies in human SLE*. Eur J Hum Genet. 2007^b; 15: 823-830.
- Graham RR, Ortmann WA, Langefeld CD, Jawaheer D, Selby SA, Rodine PR *et al*. *Visualizing human leukocyte antigen class II risk haplotypes in human systemic lupus erythematosus*. Am J Hum Genet. 2002; 71: 543-553.
- Gray-McGuire C, Moser KL, Gaffney PM, Kelly J, Yu H, Olson JM *et al*. *Genome scan of human systemic lupus erythematosus by regression modeling: evidence of linkage and epistasis at 4p16-15.2*. Am J Hum Genet. 2000; 67: 1460-1469.
- Gregersen PK, Amos CI, Lee AT, Lu Y, Remmers EF, Kastner DL *et al*. *REL, encoding a member of the NF-kappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis*. Nat Genet. 2009; 41: 820-823.
- Gregersen PK, Behrens TW. *Genetics of autoimmune diseases: disorders of immune homeostasis*. Nat Rev Genet. 2006^a; 7: 917-928.
- Gregersen PK, Lee HS, Batliwalla F, Begovich AB. *PTPN22: setting thresholds for autoimmunity*. Semin Immunol. 2006^b; 18: 214-223.
- Griffiths MM, Encinas JA, Remmers EF, Kuchroo VK, Wilder RL. *Mapping autoimmunity genes*. Curr Opin Immunol. 1999; 11: 689-700.
- Guillemin F. *Primer: the fallacy of subgroup analysis*. Nat Clin Pract Rheumatol. 2007; 3: 407-413.

- Guo L, Deshmukh H, R Lu R, Vidal GS, Kelly JA, Kaufman KM *et al.* *Replication of the BANK1 genetic association with systemic lupus erythematosus in a European-derived population.* Genes Immun. 2009; 10: 531-538.
- Guyot B, Mouchiroud G. *Characterization of promoter elements directing Mona/Gads molecular adapter expression in T and myelomonocytic cells: involvement of the AML-1 transcription factor.* J Leukoc Biol. 2003; 73: 263-272.
- Hallberg B, Thornell A, Holm M, Grundstrom T. *SEFI binding is important for T cell specific enhancers of genes for T cell receptor-CD3 subunits.* Nucleic Acids Res. 1992; 20: 6495-6499.
- Han JW, Zheng HF, Cui Y, Sun LD, Ye DQ, Hu Z *et al.* *Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus.* Nat Genet. 2009; 41: 1234-1237.
- Han S, Kim-Howard X, Deshmukh H, Kamatani Y, Viswanathan P, Guthridge JM *et al.* *Evaluation of imputation-based association in and around the integrin-alpha-M (ITGAM) gene and replication of robust association between a non-synonymous functional variant within ITGAM and systemic lupus erythematosus (SLE).* Hum Mol Genet. 2009; 18: 1171-1180.
- Hellquist A, Järvinen TM, Koskenmies S, Zucchelli M, Orsmark-Pietras C, Berglind L *et al.* *Evidence for genetic association and interaction between the TYK2 and IRF5 genes in systemic lupus erythematosus.* J Rheumatol. 2009; 36: 1631-1638.
- Hernandez-Munain C, Krangel MS. *c-Myb and core-binding factor/PEBP2 display functional synergy but bind independently to adjacent sites in the T-cell receptor delta enhancer.* Mol Cell Biol. 1995; 15: 3090-3099.
- Heyninck K, Kreike MM, Beyaert R. *Structure-function analysis of the A20-binding inhibitor of NF- κ B activation, ABIN-1.* FEBS Lett. 2003; 536: 135-140.

- Hildner KM, Schirmacher P, Atreya I, Dittmayer M, Bartsch B, Galle PR *et al.* *Targeting of the transcription factor STAT4 by antisense phosphorothioate oligonucleotides suppresses collagen-induced arthritis.* J Immunol. 2007; 178 (6): 3427-3436.
- Hinks A, Barton A, John S, Bruce I, Hawkins C, Griffiths CE *et al.* *Association between the PTPN22 gene and rheumatoid arthritis and juvenile idiopathic arthritis in a UK population: further support that PTPN22 is an autoimmunity gene.* Arthritis Rheum. 2005; 52: 1694-1699.
- Hinks A, Worthington J, Thomson W. *The association of PTPN22 with rheumatoid arthritis and juvenile idiopathic arthritis.* Rheumatology (Oxford). 2006; 45: 365-368.
- Hochberg MC. *Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus.* Arthritis Rheum. 1997; 40: 1725.
- Hollenhorst PC, Shah AA, Hopkins C, Graves BJ. *Genomewide analyses reveal properties of redundant and specific promoter occupancy within the ETS gene family.* Genes Dev. 2007; 21: 1882-1894.
- Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S *et al.* *Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX.* N Engl J Med. 2008; 358: 900-909.
- Hulley SB. *Designing clinical research.* 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2007.
- International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL *et al.* *A second generation human haplotype map of over 3.1 million SNPs.* Nature. 2007; 449: 851-861.
- Jacob CO, Reiff A, Armstrong DL, Myones BL, Silverman E, Klein-Gitelman M *et al.* *Identification of novel susceptibility genes in childhood-onset systemic*

lupus erythematosus using a uniquely designed candidate gene pathway platform. Arthritis Rheum. 2007; 56: 4164-4173.

- Jacob CO, Zang S, Li L, Ciobanu V, Quismorio F, Mizutani A *et al.* *Pivotal role of Stat4 and Stat6 in the pathogenesis of the lupus-like disease in the New Zealand mixed 2328 mice.* J Immunol. 2003; 171: 1564-1571.
- Jacob CO, Zhu J, Armstrong DL, Yan M, Han J, Zhou XJ *et al.* *Identification of IRAK1 as a risk gene with critical role in the pathogenesis of systemic lupus erythematosus.* Proc Natl Acad Sci U S A. 2009; 106: 6256–6261.
- James ES, Harney S, Wordsworth BP, Cookson WO, Davis SJ, Moffatt MF *et al.* *PDCD1: a tissue-specific susceptibility locus for inherited inflammatory disorders.* Genes Immun. 2005; 6: 430-437.
- Jawaheer D, Seldin MF, Amos CI, Chen WV, Shigeta R, Etzel C *et al.* *Screening the genome for rheumatoid arthritis susceptibility genes: a replication study and combined analysis of 512 multicase families.* Arthritis Rheum. 2003; 48: 906-916.
- Ji JD, Lee WJ, Kong KA, Woo JH, Choi SJ, Lee YH *et al.* *Association of STAT4 polymorphism with rheumatoid arthritis and systemic lupus erythematosus: a meta-analysis.* Mol Biol Rep. 2010; 37: 141-147.
- Johansson AC, Lindqvist AK, Johannesson M, Holmdahl R. *Genetic heterogeneity of autoimmune disorders in the nonobese diabetic mouse.* Scand J Immunol. 2003; 57: 203-213.
- Johansson CM, Zunec R, Garcia MA, Scherbarth HR, Tate GA, Paira S *et al.* *Chromosome 17p12-q11 harbors susceptibility loci for systemic lupus erythematosus.* Hum Genet. 2004; 115:230-238.
- Johansson M, Arlestig L, Möller B, Rantapää-Dahlqvist S. *Association of a PDCD1 polymorphism with renal manifestations in systemic lupus erythematosus.* Arthritis Rheum. 2005; 52: 1665-1669.

- Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N *et al.* *Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription.* Nat Genet. 1998; 19: 187-191.
- Kamachi Y, Ogawa E, Asano M, Ishida S, Murakami Y, Satake M *et al.* *Purification of a mouse nuclear factor that binds to both the A and B cores of the polyomavirus enhancer.* J Virol. 1990; 64: 4808-4819.
- Karassa FB, Trikalinos TA, Ioannidis JP. *The Fc gamma RIIIA-F158 allele is a risk factor for the development of lupus nephritis: a meta-analysis.* Kidney Int. 2003; 63: 1475-1482.
- Kariuki SN, Kirou KA, MacDermott EJ, Barillas-Arias L, Crow MK, Niewold TB. *Cutting edge: autoimmune disease risk variant of STAT4 confers increased sensitivity to IFN-alpha in lupus patients in vivo.* J Immunol. 2009; 182: 34-38.
- Kassel R, Cruise MW, Iezzoni JC, Taylor NA, Pruett TL, Hahn YS. *Chronically inflamed livers up-regulate expression of inhibitory B7 family members.* Hepatology. 2009; 50: 1625-1637.
- Kaufman KM, Kelly JA, Herring BJ, Adler AJ, Glenn SB, Namjou B *et al.* *Evaluation of the Genetic Association of the PTPN22 R620W Polymorphism in Familial and Sporadic Systemic Lupus Erythematosus.* Arthritis Rheum. 2006; 54: 2533-2540.
- Kawasaki A, Ito I, Hikami K, Ohashi J, Hayashi T, Goto D *et al.* *Role of STAT4 polymorphisms in systemic lupus erythematosus in a Japanese population: a case-control association study of the STAT1-STAT4 region.* Arthritis Res Ther. 2008^a; 10: R113.
- Kawasaki A, Kyogoku C, Ohashi J, Miyashita R, Hikami K, Kusaoi M *et al.* *Association of IRF5 polymorphisms with systemic lupus erythematosus in a Japanese population: support for a crucial role of intron 1 polymorphisms.* Arthritis Rheum. 2008^b; 58: 826-834.

- Keir ME, Francisco LM, Sharpe AH. *PD-1 and its ligands in T-cell immunity*. Curr Opin Immunol. 2007; 19: 309-314.
- Kel AE, Gossling E, Reuter I, Cheremushkin E, Kel-Margoulis OV, Wingender E *et al*. *MATCH: A tool for searching transcription factor binding sites in DNA sequences*. Nucleic Acids Res. 2003; 31: 3576-3579.
- Kelly JA, Kelley JM, Kaufman KM, Kilpatrick J, Bruner GR, Merrill JT *et al*. *Interferon regulatory factor-5 is genetically associated with systemic lupus erythematosus in African Americans*. Genes Immun. 2008; 9: 187-194.
- Kelly JA, Moser KL, Harley JB. *The genetics of systemic lupus erythematosus: putting the pieces together*. Genes Immun. 2002; 3 Suppl 1:S71-S85.
- Kim-Howard X, Maiti AK, Anaya JM, Bruner GR, Brown E, Merrill JT *et al*. *ITGAM coding variant (rs1143679) influences the risk of renal disease, discoid rash, and immunologic manifestations in lupus patients with European ancestry*. Ann Rheum Dis. 2009. [Epub ahead of print].
- Klinman DM, Gursel I, Klaschik S, Dong L, Currie D, Shirota H. *Therapeutic potential of oligonucleotides expressing immunosuppressive TTAGGG motifs*. Ann N Y Acad Sci. 2005; 1058: 87-95.
- Klose RJ, Bird AP. *Genomic DNA methylation: the mark and its mediators*. Trends Biochem Sci. 2006; 31: 89-97.
- Kobayashi S, Ikari K, Kaneko H, Kochi Y, Yamamoto K, Shimane K *et al*. *Association of STAT4 with susceptibility to rheumatoid arthritis and systemic lupus erythematosus in the Japanese population*. Arthritis Rheum. 2008; 58: 1940-1946.
- Korman BD, Alba MI, Le JM, Alevizos I, Smith JA, Nikolov NP *et al*. *Variant form of STAT4 is associated with primary Sjogren's syndrome*. Genes Immun. 2008; 9:267-270.

- Koskenmies S, Lahermo P, Julkunen H, Ollikainen V, Kere J, Widen E. *Linkage mapping of systemic lupus erythematosus (SLE) in Finnish families multiply affected by SLE*. J Med Genet. 2004; 41: e2-e5.
- Kotenko SV, Pestka S. *Jak-Stat signal transduction pathway through the eyes of cytokine class II receptor complexes*. Oncogene. 2000; 19: 2557-2565.
- Kozyrev SV, Abelson AK, Wojcik J, Zaghlool A, Linga Reddy MVP, Sanchez E *et al*. *Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus*. Nat Genet. 2008. 40 (2): 211-216.
- Kozyrev SV, Lewen S, Reddy PM, Pons-Estel B, Witte T, Junker P *et al*. *Structural insertion/deletion variation in IRF5 is associated with a risk haplotype and defines the precise IRF5 isoforms expressed in systemic lupus erythematosus*. Arthritis Rheum. 2007; 56: 1234-1241.
- Kyogoku C, Langefeld CD, Ortmann WA, Lee A, Selby S, Carlton VE *et al*. *Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE*. Am J Hum Genet. 2004; 75: 504-507.
- Latchman Y, Wood CR, Chernova T, Chaudhary D, Borde M, Chernova I *et al*. *PD-L2 is a second ligand for PD-1 and inhibits T cell activation*. Nat Immunol. 2001; 2: 261-268.
- Lauwerys BR, Wakeland EK. *Genetics of lupus nephritis*. Lupus. 2005; 14: 2-12.
- Lawrence JS, Martins CL, Drake GL. *A family survey of lupus erythematosus. I. Heritability*. J Rheumatol. 1987; 14: 913-921.
- Lee EG, Boone DL, Chai S, Libby SL, Chien M, Lodolce JP *et al*. *Failure to regulate TNF-induced NF-kappaB and cell death responses in A20-deficient mice*. Science. 2000; 289: 2350-2354.
- Lee HS, Remmers EF, Le JM, Kastner DL, Bae SC, Gregersen PK. *Association of STAT4 with rheumatoid arthritis in the Korean population*. Mol Med. 2007; 13: 455-460.

- Lee YH, Nath SK. *Systemic lupus erythematosus susceptibility loci defined by genome scan meta-analysis*. Hum Genet. 2005; 118: 434-443.
- Lee-Kirsch MA, Gong M, Chowdhury D, Senenko L, Engel K, Lee YA *et al*. *Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 are associated with systemic lupus erythematosus*. Nat Genet. 2007; 39: 1065-1067.
- Leibson PJ. *The regulation of lymphocyte activation by inhibitory receptors*. Curr Opin Immunol. 2004; 16: 328-336.
- Li CC, Ramirez-Carrozzi VR, Smale ST. *Pursuing gene regulation 'logic' via RNA interference and chromatin immunoprecipitation*. Nat Immunol. 2006; 7: 692-697.
- Li Q, Johnston SA. *Are all DNA binding and transcription regulation by an activator physiologically relevant?* Mol Cell Biol. 2001; 21: 2467-2474.
- Li QJ, Chau J, Ebert PJ, Sylvester G, Min H, Liu G *et al*. *miR-181a is an intrinsic modulator of T cell sensitivity and selection*. Cell. 2007; 129: 147-161.
- Lindqvist AK, Steinsson K, Johanneson B, Kristjansdottir H, Arnasson A, Grondal G *et al*. *A susceptibility locus for human systemic lupus erythematosus (hSLE1) on chromosome 2q*. J Autoimmun. 2000; 14: 169-178.
- MacKenzie A, Quinn J. *A serotonin transporter gene intron 2 polymorphic region, correlated with affective disorders, has allele-dependent differential enhancer-like properties in the mouse embryo*. Proc Natl Acad Sci U S A. 1999; 96: 15251-15255.
- Magnusson V, Lindqvist AK, Castillejo-Lopez C, Kristjansdottir H, Steinsson K, Grondal G *et al*. *Fine mapping of the SLEB2 locus involved in susceptibility to systemic lupus erythematosus*. Genomics. 2000; 70: 307-314.
- Manderson AP, Botto M, Walport MJ. *The role of complement in the development of systemic lupus erythematosus*. Annu Rev Immunol. 2004; 22: 431-456.

- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ *et al.* *Finding the missing heritability of complex diseases.* Nature. 2009; 461: 747-753.
- Martinez A, Varade J, Marquez A, Cenit MC, Espino L, Perdigonés N *et al.* *Association of the STAT4 gene with increased susceptibility for some immune-mediated diseases.* Arthritis Rheum. 2008; 58: 2598-2602.
- Mevorach D, Mascarenhas JO, Gershov D, Elkon KB. *Complement-dependent clearance of apoptotic cells by human macrophages.* J Exp Med. 1998; 188: 2313-2320.
- Meyers S, Downing JR, Hiebert SW. *Identification of AML-1 and the (8;21) translocation protein (AML-1/ETO) as sequence-specific DNA-binding proteins: the runt homology domain is required for DNA binding and protein-protein interactions.* Mol Cell Biol. 1993; 13: 6336-6345.
- Miller W, Rosenbloom K, Hardison RC, Hou M, Taylor J, Raney B *et al.* *28-way vertebrate alignment and conservation track in the UCSC Genome Browser.* Genome Res. 2007; 17: 1797-1808.
- Minegishi Y, Saito M, Morio T, Watanabe K, Agematsu K, Tsuchiya S *et al.* *Human tyrosine kinase 2 deficiency reveals its requisite roles in multiple cytokine signals involved in innate and acquired immunity.* Immunity 2006; 25: 745-755.
- Mishra N, Reilly CM, Brown DR, Ruiz P, Gilkeson GS. *Histone deacetylase inhibitors modulate renal disease in the MRL- lpr/ lpr mouse.* J Clin Invest. 2003; 111: 539-552.
- Mok CC, Lau CS. *Pathogenesis of systemic lupus erythematosus.* J Clin Pathol. 2003; 56: 481-490.
- Moser KL, Kelly JA, Lessard CJ, Harley JB. *Recent insights into the genetic basis of systemic lupus erythematosus.* Genes Immun. 2009; 10: 373-379.

- Moser KL, Neas BR, Salmon JE, Yu H, Gray-McGuire C, Asundi N *et al.* *Genome scan of human systemic lupus erythematosus: evidence for linkage on chromosome 1q in African-American pedigrees.* Proc Natl Acad Sci U S A. 1998; 95: 14869-14874.
- Moyé LA. *Multiple analyses in clinical trials: fundamentals for investigators.* New York: Springer; 2003.
- Musone SL, Taylor KE, Lu TT, Nititham J, Ferreira RC, Ortmann W *et al.* *Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus.* Nat Genet. 2008; 40: 1062-1064.
- Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, Goldgar D *et al.* *Genome-wide scan reveals association of psoriasis with IL-23 and NF- κ B pathways.* Nat Genet. 2009; 41: 199-204.
- Nalbandian A, Crispín JC, Tsokos GC. *Interleukin-17 and systemic lupus erythematosus: current concepts.* Clin Exp Immunol. 2009; 157: 209-215
- Namjou B, Sestak AL, Armstrong DL, Zidovetzki R, Kelly JA, Jacob N *et al.* *High-density genotyping of STAT4 reveals multiple haplotypic associations with systemic lupus erythematosus in different racial groups.* Arthritis Rheum. 2009; 60: 1085-1095.
- Nath SK, Han S, Kim-Howard X, Kelly JA, Viswanathan P, Gilkeson GS *et al.* *A nonsynonymous functional variant in integrin- α M (encoded by ITGAM) is associated with systemic lupus erythematosus.* Nat Genet. 2008. 40: 152-154.
- Nath SK, Harley JB, Lee YH. *Polymorphisms of complement receptor 1 and interleukin-10 genes and systemic lupus erythematosus: a meta-analysis.* Hum Genet. 2005; 118: 225-234.
- Nath SK, Quintero-Del-Rio AI, Kilpatrick J, Feo L, Ballesteros M, Harley JB. *Linkage at 12q24 with systemic lupus erythematosus (SLE) is established and confirmed in Hispanic and European American families.* Am J Hum Genet. 2004; 74: 73-82.

- Nepom GT, Byers P, Seyfried C, Healey LA, Wilske KR, Stage D *et al.* *HLA genes associated with rheumatoid arthritis: identification of susceptibility alleles using specific oligonucleotide probes.* Arthritis Rheum. 1989; 32: 15-21.
- Newby PR, Roberts-Davies EL, Brand OJ, Heward JM, Franklyn JA, Gough SC *et al.* *Tag SNP screening of the PDCD1 gene for association with Graves' disease.* Clin Endocrinol (Oxf). 2007; 67: 125-128.
- Nielsen C, Hansen D, Husby S, Jacobsen BB, Lillevang ST. *Association of a putative regulatory polymorphism in the PD-1 gene with susceptibility to type 1 diabetes.* Tissue Antigens. 2003; 62: 492-497.
- Nielsen C, Laustrop H, Voss A, Junker P, Husby S, Lillevang ST. *A putative regulatory polymorphism in PD-1 is associated with nephropathy in a population-based cohort of systemic lupus erythematosus patients.* Lupus. 2004; 13: 510-516.
- Niewold TB, Kelly JA, Flesch MH, Espinoza LR, Harley JB and Crow MK. *Association of the IRF5 risk haplotype with high serum interferon-alpha activity in systemic lupus erythematosus patients.* Arthritis Rheum. 2008; 58 (8): 2481-2487.
- Nishikomori R, Usui T, Wu CY, Morinobu A, O'Shea JJ, Strober W. *Activated STAT4 has an essential role in Th1 differentiation and proliferation that is independent of its role in the maintenance of IL-12R beta 2 chain expression and signaling.* J Immunol. 2002; 169: 4388-4398.
- Nishimura H, Minato N, Nakano T, Honjo T. *Immunological studies on PD-1 deficient mice: implication of PD-1 as a negative regulator for B cell responses.* Int Immunol. 1998; 10: 1563-1572.
- Nishimura H, Nose M, Hiai H, Minato N, Honjo T. *Development of lupus-like autoimmune diseases by disruption of the PD-1 gene encoding an ITIM motif-carrying immunoreceptor.* Immunity. 1999; 11: 141-151.

- Nordmark G, Kristjansdottir G, Theander E, Eriksson P, Brun JG, Wang C *et al.* *Additive effects of the major risk alleles of IRF5 and STAT4 in primary Sjögren's syndrome.* Genes Immun. 2009; 10: 68-76.
- Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A *et al.* *Genes mirror geography within Europe.* Nature. 2008; 456: 98-101.
- Oishi Y, Sumida T, Sakamoto A, Kita Y, Kurasawa K, Nawata Y *et al.* *Selective reduction and recovery of invariant Valpha24JalphaQ T cell receptor T cells in correlation with disease activity in patients with systemic lupus erythematosus.* J Rheumatol. 2001; 28: 275-283.
- Okazaki T, Tanaka Y, Nishio R, Mitsuiye T, Mizoguchi A, Wang J *et al.* *Autoantibodies against cardiac troponin I are responsible for dilated cardiomyopathy in PD-1-deficient mice.* Nat Med. 2003; 9: 1477-1483.
- Olson JM, Song Y, Dudek DM, Moser KL, Kelly JA, Bruner GR *et al.* *A genome screen of systemic lupus erythematosus using affected-relative-pair linkage analysis with covariates demonstrates genetic heterogeneity.* Genes Immun. 2002; 3 Suppl 1: S5-S12.
- Onengut-Gumuscu S, Ewens KG, Spielman RS, Concannon P. *A functional polymorphism (1858C/T) in the PTPN22 gene is linked and associated with type I diabetes in multiplex families.* Genes Immun. 2004; 5: 678-80.
- Ono M, Yaguchi H, Ohkura N, Kitabayashi I, Nagamura Y, Nomura T *et al.* *Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1.* Nature. 2007; 446: 685-689.
- Orozco G, Abelson AK, Gonzalez-Gay MA, Balsa A, Pascual-Salcedo D, Garcia A *et al.* *Study of functional variants of the BANK1 gene in rheumatoid arthritis.* Arthritis Rheum. 2009^a; 60: 372-379.
- Orozco G, Alizadeh BZ, Delgado-Vega AM, Gonzalez-Gay MA, Balsa A, Pascual-Salcedo D *et al.* *Association of STAT4 with rheumatoid arthritis: a*

replication study in three European populations. Arthritis Rheum. 2008; 58:1974-1980.

- Orozco G, Hinks A, Eyre S, Ke X, Gibbons LJ, Bowes J et al. *Combined effects of three independent SNPs greatly increase the risk estimate for RA at 6q23.* Hum Mol Genet. 2009^b; 18: 2693-2699.
- Orozco G, Sanchez E, Gonzalez-Gay MA, Lopez-Nevot MA, Torres B, Caliz R et al. *Association of a functional single-nucleotide polymorphism of PTPN22, encoding lymphoid protein phosphatase, with rheumatoid arthritis and systemic lupus erythematosus.* Arthritis Rheum. 2005; 52: 219-224.
- Orrú V, Tsai SJ, Rueda B, Fiorillo E, Stanford SM, Dasgupta J et al. *A loss-of-function variant of PTPN22 is associated with reduced risk of systemic lupus erythematosus.* Hum Mol Genet. 2009; 18: 569-579.
- Palomino-Morales RJ, Rojas-Villarraga A, Gonzalez CI, Ramírez G, Anaya JM, Martín J. *STAT4 but not TRAF1/C5 variants influence the risk of developing rheumatoid arthritis and systemic lupus erythematosus in Colombians.* Genes Immun. 2008; 9: 379-382.
- Panush RS, Edwards LN, Longley S, Webster E. *"Rheupus" syndrome.* Arch Intern Med. 1988; 148: 1633-1636.
- Peng G, Li S, Wu W, Tan X, Chenb Y, Chen Z. *PD-1 upregulation is associated with HBV-specific T cell dysfunction in chronic hepatitis B patients.* Mol Immunol. 2008; 45: 963-970.
- Peschken CA, Katz SJ, Silverman E, Pope JE, Fortin PR, Pineau C et al. *The 1000 Canadian faces of lupus: determinants of disease outcome in a large multiethnic cohort.* J Rheumatol. 2009; 36: 1200-1208.
- Petri M. *Epidemiology of systemic lupus erythematosus.* Best Pract Res Clin Rheumatol. 2002; 16: 547-558.
- Petri M. *The effect of race on the presentation and course of SLE in the United States.* Arthritis Rheum. 1997; 40: 785-785.

- Petrovas C, Casazza JP, Brenchley JM, Price DA, Gostick E, Adams WC *et al.* *PD-1 is a regulator of virus-specific CD8⁺ T cell survival in HIV infection.* J Exp Med. 2006; 203 (10): 2281-2292
- Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PIW, Maller J *et al.* *Two independent alleles at 6q23 associated with risk of rheumatoid arthritis.* Nat Genet. 2007^a; 39: 1477-1482.
- Plenge RM, Padyukov L, Remmers EF, Purcell S, Lee AT, Karlson EW *et al.* *Replication of putative candidate-gene associations with rheumatoid arthritis in 4,000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4.* Am J Hum Genet. 2005; 77: 1044-1060.
- Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B *et al.* *TRAF1-C5 as a risk locus for rheumatoid arthritis: a genomewide study.* N Engl J Med. 2007^b; 357: 1199-1209.
- Pons-Estel BA, Catoggio LJ, Cardiel MH, Soriano ER, Gentiletti S, Villa AR *et al.* *The GLADEL multinational Latin American prospective inception cohort of 1,214 patients with systemic lupus erythematosus: ethnic and disease heterogeneity among "Hispanics".* Medicine (Baltimore). 2004; 83: 1-17.
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. *Principal components analysis corrects for stratification in genome-wide association studies.* Nat Genet. 2006; 38: 904-909.
- Prokunina L, Castillejo-Lopez C, Oberg F, Gunnarsson I, Berg L, Magnusson V *et al.* *A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans.* Nat Genet. 2002; 32: 666-669.
- Prokunina L, Gunnarsson I, Sturfelt G, Truedsson L, Seligman VA, Olson JL *et al.* *The systemic lupus erythematosus-associated PDCD1 polymorphism PD1.3A in lupus nephritis.* Arthritis Rheum. 2004; 50: 327-328.
- Radziejewicz H, Dunham RM, Grakoui A. *PD-1 tempers Tregs in chronic HCV infection.* J Clin Invest. 2009; 119: 450-453.

- Ragimbeau J, Dondi E, Alcover A, Eid P, Uze G, Pellegrini S. *The tyrosine kinase Tyk2 controls IFNAR1 cell surface expression*. EMBO J. 2003; 22: 537-547.
- Rahmann S, Muller T, Vingron M. *On the power of profiles for transcription factor binding site detection*. Stat Appl Genet Mol Biol. 2003; 2: Article7.
- Raman K, Mohan C. *Genetic underpinnings of autoimmunity: lessons from studies in arthritis, diabetes, lupus and multiple sclerosis*. Curr Opin Immunol. 2003; 15: 651-659.
- Ramos PS, Kelly JA, Gray-McGuire C, Bruner GR, Leiran AN, Meyer CM *et al*. *Familial aggregation and linkage analysis of autoantibody traits in pedigrees multiplex for systemic lupus erythematosus*. Genes Immun. 2006; 7: 417-432.
- Rao S, Olson JM, Moser KL, Gray-McGuire C, Bruner GR, Kelly J *et al*. *Linkage analysis of human systemic lupus erythematosus-related traits: a principal component approach*. Arthritis Rheum. 2001; 44: 2807-2818.
- Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Guiducci C, Burt NP *et al*. *Common variants at CD40 and other loci confer risk of rheumatoid arthritis*. Nat Genet. 2008; 40: 1216-1223.
- Reddy MV, Johansson M, Sturfelt G, Jonsen A, Gunnarsson I, Svenungsson E, *et al*. *The R620W C/T polymorphism of the gene PTPN22 is associated with SLE independently of the association of PDCDI*. Genes Immun. 2005; 6: 658-662.
- Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD *et al*. *Global variation in copy number in the human genome*. Nature. 2006; 444: 444-454.
- Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW *et al*. *STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus*. N Engl J Med. 2007; 357: 977-986.

- Ren Y, Tang J, Mok MY, Chan AW, Wu A, Lau CS. *Increased apoptotic neutrophils and macrophages and impaired macrophage phagocytic clearance of apoptotic neutrophils in systemic lupus erythematosus*. *Arthritis Rheum*. 2003; 48: 2888-2897.
- Reveille JD, Arnett FC, Wilson RW, Bias WB, McLean RH. *Null alleles of the fourth component of complement and HLA haplotypes in familial systemic lupus erythematosus*. *Immunogenetics*. 1985; 21: 299-311.
- Richardson B, Scheinbart L, Strahler J, Gross L, Hanash S, Johnson M. *Evidence for impaired T cell DNA methylation in systemic lupus erythematosus and rheumatoid arthritis*. *Arthritis Rheum*. 1990; 33: 1665-1673.
- Richman IB, Chung SA, Taylor KE, Kosoy R, Tian C, Ortmann WA *et al*. *European population substructure correlates with systemic lupus erythematosus endophenotypes in North Americans of European descent*. *Genes Immun*. 2009. [Epub ahead of print].
- Richter MF, Dumenil G, Uze G, Fellous M, Pellegrini S. *Specific contribution of TYK2 JH regions to the binding and the expression of the interferon α/β receptor component IFNAR1*. *J Biol Chem*. 1998; 273: 24723-24729.
- Rieck M, Arechiga A, Onengut-Gumuscu S, Greenbaum C, Concannon P, Buckner JH. *Genetic variation in PTPN22 corresponds to altered function of T and B lymphocytes*. *J Immunol*. 2007; 179: 4704-4710.
- Roulet E, Fisch I, Junier T, Bucher P, Mermod N. *Evaluation of computer tools for the prediction of transcription factor binding sites on genomic DNA*. In *Silico Biol*. 1998; 1: 21-28.
- Salloum R, Franek BS, Kariuki SN, Rhee L, Mikolaitis RA, Jolly M *et al*. *Genetic variation at the IRF7/PHRF1 locus is associated with autoantibody profile and serum interferon-alpha activity in lupus patients*. *Arthritis Rheum*. 2010; 62: 553-561.

- Sanghera DK, Manzi S, Bontempo F, Nestlerode C, Kamboh MI. *Role of an intronic polymorphism in the PDCD1 gene with the risk of sporadic systemic lupus erythematosus and the occurrence of antiphospholipid antibodies.* Hum Genet. 2004; 115: 393-398.
- Sawalha AH, Webb R, Han S, Kelly JA, Kaufman KM, Kimberly RP *et al.* *Common variants within MECP2 confer risk of systemic lupus erythematosus.* PLoS ONE. 2008; 3: e1727.
- Scofield RH, Bruner GR, Kelly JA, Kilpatrick J, Bacino D, Nath SK *et al.* *Thrombocytopenia identifies a severe familial phenotype of systemic lupus erythematosus and reveals genetic linkages at 1q22 and 11p13.* Blood. 2003; 101: 992-997.
- Sebastiani GD, Galeazzi M. *Immunogenetic studies on systemic lupus erythematosus.* Lupus. 2009; 18: 878-883.
- Seldin MF, Amos CI, Ward R, Gregersen PK. *The genetics revolution and the assault on rheumatoid arthritis.* Arthritis Rheum. 1999; 42: 1071-1079.
- Seldin MF, Shigeta R, Villoslada P, Selmi C, Tuomilehto J, Silva G *et al.* *European population substructure: Clustering of northern and southern populations.* PLoS Genet. 2006; 2: e143.
- Serrano NC, Millan P, Páez MC. *Non-HLA associations with autoimmune diseases.* Autoimmun Rev. 2006; 5: 209-214.
- Sestak AL, Nath SK, Kelly JA, Bruner GR, James JA, Harley JB. *Patients with familial and sporadic onset SLE have similar clinical profiles but vary profoundly by race.* Lupus. 2008; 17: 1004-1009.
- Sestak AL, Nath SK, Sawalha AH, Harley JB. *Current status of lupus genetics.* Arthritis Res Ther. 2007; 9: 210.
- Shai R, Quismorio FP Jr, Li L, Kwon OJ, Morrison J, Wallace DJ *et al.* *Genome-wide screen for systemic lupus erythematosus susceptibility genes in multiplex families.* Hum Mol Genet. 1999; 8: 639-644.

- Shappell SB, Toman C, Anderson DC, Taylor AA, Entman ML, Smith CW. *Mac-1 (CD11b/CD18) mediates adherence-dependent hydrogen peroxide production by human and canine neutrophils.* J Immunol. 1990; 144: 2702-2711.
- Sharpe AH, Freeman GJ. *The B7-CD28 superfamily.* Nat Rev Immunol. 2002; 2: 116-126.
- Shimane K, Kochi Y, Horita T, Ikari K, Amano H, Hirakata M *et al.* *The association of a nonsynonymous single-nucleotide polymorphism in TNFAIP3 with systemic lupus erythematosus and rheumatoid arthritis in the Japanese population.* Arthritis Rheum. 2010; 62: 574-579.
- Shin HD, Kim I, Choi CB, Lee SO, Lee HW, Bae SC. *Different genetic effects of interferon regulatory factor 5 (IRF5) polymorphisms on systemic lupus erythematosus in a Korean population.* J Rheumatol. 2008; 35: 2148-2151.
- Sigurdsson S, Göring HH, Kristjansdottir G, Milani L, Nordmark G, Sandling JK *et al.* *Comprehensive evaluation of the genetic variants of interferon regulatory factor 5 (IRF5) reveals a novel 5 bp length polymorphism as strong risk factor for systemic lupus erythematosus.* Hum Mol Genet. 2008^a; 17: 872-881.
- Sigurdsson S, Nordmark G, Garnier S, Grundberg E, Kwan T, Nilsson O *et al.* *A risk haplotype of STAT4 for systemic lupus erythematosus is over-expressed, correlates with antidsDNA and shows additive effects with two risk alleles of IRF5.* Hum Mol Genet. 2008^b; 17: 2868-2876.
- Sigurdsson S, Nordmark G, Göring HH, Lindroos K, Wiman AC, Sturfelt G, *et al.* *Polymorphisms in the Tyrosine Kinase 2 and Interferon Regulatory Factor 5 Genes Are Associated with Systemic Lupus Erythematosus.* Am J Hum Genet. 2005. 76: 528-537.

- Sigurdsson S, Padyukov L, Kurreeman FA, Liljedahl U, Wiman AC, Alfredsson L *et al.* *Association of a haplotype in the promoter region of the interferon regulatory factor 5 gene with rheumatoid arthritis.* Arthritis Rheum. 2007; 56: 2202-2210.
- Silman AJ, Macgregor AJ, Thomson W, Holligan S, Carthy D, Farhan A, Ollier WE. *Twin concordance rates for rheumatoid arthritis: Results from a nationwide study.* Br J Rheumatol. 1993; 32: 903-907.
- Singh RR, Ebling FM, Albuquerque DA, Saxena V, Kumar V, Giannini EH *et al.* *Induction of autoantibody production is limited in nonautoimmune mice.* J Immunol. 2002; 169: 587-594.
- Singh RR, Saxena V, Zang S, Li L, Finkelman FD, Witte DP *et al.* *Differential contribution of IL-4 and STAT6 vs STAT4 to the development of lupus nephritis.* J Immunol. 2003; 170: 4818-4825.
- SLEGEN, Harley JB, Alarcón-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP *et al.* *Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXX, KIAA1542 and other loci.* Nat Genet. 2008; 40: 204-210.
- Sohn JH, Bora PS, Suk HJ, Molina H, Kaplan HJ, Bora NS. *Tolerance is dependent on complement C3 fragment iC3b binding to antigen-presenting cells.* Nat Med. 2003; 9: 206-212.
- Stetson DB, Ko JS, Heidmann T, Medzhitov R. *Trex1 prevents cell-intrinsic initiation of autoimmunity.* Cell. 2008; 134: 587-598.
- Stohl W, Metyas S, Tan SM, Cheema GS, Oamar B, Xu D *et al.* *B lymphocyte stimulator overexpression in patients with systemic lupus erythematosus: longitudinal observations.* Arthritis Rheum. 2003; 48: 3475-3486.
- Stormo GD. *DNA binding sites: representation and discovery.* Bioinformatics. 2000; 16: 16-23.

- Su K, Wu J, Edberg JC, Li X, Ferguson P, Cooper GS *et al.* *A promoter haplotype of the immunoreceptor tyrosine-based inhibitory motif-bearing FcγRIIb alters receptor expression and associates with autoimmunity. I. Regulatory FCGR2B polymorphisms and their association with systemic lupus erythematosus.* J Immunol. 2004; 172: 7186-7191.
- Sullivan KE, Petri MA, Schmeckpeper BJ, McLean RH, Winkelstein JA. *Prevalence of a mutation causing C2 deficiency in systemic lupus erythematosus.* J Rheumatol. 1994; 21: 1128-1133.
- Svenungsson E, Gustafsson J, Leonard D, Sandling J, Gunnarsson I, Nordmark G *et al.* *A STAT4 risk allele is associated with ischemic cerebrovascular events and antiphospholipid antibodies in systemic lupus erythematosus.* Ann Rheum Dis. 2009. [Epub ahead of print].
- Taganov KD, Boldin MP, Chang KJ, Baltimore D. *NFκB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses.* Proc Natl Acad Sci U S A. 2006; 103: 12481-12486.
- Takeuchi H, Takeuchi T, Gao J, Cantley LC, Hirata M. *Characterization of PDK as a protein involved in epidermal growth factor receptor trafficking.* Mol Cell Biol. 2010. [Epub ahead of print].
- Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF *et al.* *The 1982 revised criteria for the classification of systemic lupus erythematosus.* Arthritis Rheum. 1982; 25:1271-1277.
- Tang Y, Luo X, Cui H, Ni X, Yuan M, Guo Y *et al.* *MicroRNA-146a contributes to abnormal activation of the type I interferon pathway in human lupus by targeting the key signaling proteins.* Arthritis Rheum. 2009; 60: 1065-1075.
- Taylor KE, Remmers EF, Lee AT, Ortmann WA, Plenge RM, Tian C *et al.* *Specificity of the STAT4 genetic association for severe disease manifestations of systemic lupus erythematosus.* PLoS Genet. 2008; 4: e1000084.

- Teubner A, Tillmann HL, Schuppan D, Gericke G, Manns MP, Stölzel U. *Prevalence of circulating autoantibodies in healthy individuals*. Med Klin (Munich). 2002; 97: 645-649.
- Texido G, Su IH, Mecklenbräuker I, Saijo K, Malek SN, Desiderio S *et al*. *The B-cell-specific Src-family kinase Blk is dispensable for B-cell development and activation*. Mol Cell Biol. 2000; 20: 1227-1233.
- Thomas G, Jacobs KB, Yeager M, Kraft P, Wacholder S, Orr N *et al*. *Multiple loci identified in a genome-wide association study of prostate cancer*. Nat Genet. 2008; 40: 310-315.
- Thomson W, Barton A, Ke X, Eyre S, Hinks A, Bowes J *et al*. *Rheumatoid arthritis association at 6q23*. Nat Genet. 2007; 39: 1431-1433.
- Thorburn CM, Prokunina-Olsson L, Sterba KA, Lum RF, Seldin MF, Alarcon-Riquelme ME *et al*. *Association of PDCD1 genetic variation with risk and clinical manifestations of systemic lupus erythematosus in a multiethnic cohort*. Genes Immun. 2007; 8: 279-287.
- Thornell A, Hallberg B, Grundstrom T. *Binding of SL3-3 enhancer factor 1 transcriptional activators to viral and chromosomal enhancer sequences*. J Virol. 1991; 65: 42-50.
- Thornell A, Hallberg B, Grundstrom T. *Differential protein binding in lymphocytes to a sequence in the enhancer of the mouse retrovirus SL3-3*. Mol Cell Biol. 1988; 8: 1625-1637.
- Tian C, Plenge RM, Ransom M, Lee A, Villoslada P, Selmi C *et al*. *Analysis and application of European genetic substructure using 300 K SNP information*. PLoS Genet. 2008; 4: e4.
- Tiwari HK, Barnholtz-Sloan J, Wineinger N, Padilla MA, Vaughan LK, Allison DB. *Review and evaluation of methods correcting for population stratification with a focus on underlying statistical principles*. Hum Hered. 2008; 66: 67-86.

- Trautmann L, Janbazian L, Chomont N, Said EA, Gimmig S, Bessette B *et al.* *Upregulation of PD-1 expression on HIV-specific CD8+ T cells leads to reversible immune dysfunction.* Nat Med. 2006; 12: 1198-1202.
- Tretter T, Ross AE, Dordai DI, Desiderio S. *Mimicry of Pre-B Cell Receptor Signaling by Activation of the Tyrosine Kinase Blk.* J Exp Med. 2003; 198: 1863-1873.
- Tsao BP, Grossman JM, Riemekasten G, Strong N, Kalsi J, Wallace DJ *et al.* *Familiality and co-occurrence of clinical features of systemic lupus erythematosus.* Arthritis Rheum. 2002; 46: 2678-2685.
- Tsao BP. *The genetics of human systemic lupus erythematosus.* Trends Immunol. 2003; 24: 595-602.
- Tsao BP. *Update on human systemic lupus erythematosus genetics.* Curr Opin Rheumatol. 2004; 16: 513-521.
- Uchida H, Zhang J, Nimer SD. *AML1A and AML1B can transactivate the human IL-3 promoter.* J Immunol. 1997; 158: 2251-2258.
- Urbani S, Amadei B, Tola D, Massari M, Schivazappa S, Missale G *et al.* *PD-1 expression in acute hepatitis C virus (HCV) infection is associated with HCV-specific CD8 exhaustion.* J Virol. 2006; 80: 11398-11403.
- Uribe AG, McGwin G Jr, Reveille JD, Alarcon GS. *What have we learned from a 10-year experience with the LUMINA (Lupus in Minorities; Nature vs. nurture) cohort? Where are we heading?* Autoimmun Rev. 2004; 3: 321-329.
- van der Pol W, van de Winkel JG. *IgG receptor polymorphisms: risk factors for disease.* Immunogenetics. 1998; 48: 222-232.
- Vang T, Congia M, Macis MD, Musumeci L, Orru V, Zavattari P *et al.* *Autoimmune-associated lymphoid tyrosine phosphatase is a gain-of-function variant.* Nat Genet. 2005; 37: 1317-1319.

- Vasconcelos C, Carvalho C, Leal B, Pereira C, Bettencourt A, Costa PP *et al.* *HLA in Portuguese systemic lupus erythematosus patients and their relation to clinical features.* Ann N Y Acad Sci. 2009; 1173: 575-580.
- Velaga MR, Wilson V, Jennings CE, Owen CJ, Herington S, Donaldson PT *et al.* *The codon 620 tryptophan allele of the lymphoid tyrosine phosphatase (LYP) gene is a major determinant of Graves' disease.* J Clin Endocrinol Metab. 2004; 89: 5862-5865.
- Velázquez-Cruz R, Orozco L, Espinosa-Rosales F, Carreño-Manjarrez R, Solís-Vallejo E, López-Lara ND *et al.* *Association of PDCD1 polymorphisms with childhood-onset systemic lupus erythematosus.* Eur J Hum Genet. 2007; 15: 336-341.
- Viken MK, Amundsen SS, Kvien TK, Boberg KM, Gilboe IM, Lilleby V *et al.* *Association analysis of the 1858C>T polymorphism in the PTPN22 gene in juvenile idiopathic arthritis and other autoimmune diseases.* Genes Immun. 2005; 6: 271-273.
- Wallace DJ, Hahn B, Dubois EL. *Dubois' lupus erythematosus.* 7th ed. Philadelphia: Lippincott Williams & Wilkin; 2007.
- Wandstrat A, Wakeland E. *The genetics of complex autoimmune diseases: non-MHC susceptibility genes.* Nat Immunol. 2001; 2: 802-809.
- Wang F, Wang CL, Tan CT, Manivasagar M. *Systemic lupus erythematosus in Malaysia: a study of 539 patients and comparison of prevalence and disease expression in different racial and gender groups.* Lupus. 1997; 6: 248-253.
- Wang J, Yoshida T, Nakaki F, Hiai H, Okazaki T, Honjo T *et al.* *Establishment of NOD-Pdcd1^{-/-} mice as an efficient animal model of type I diabetes.* Proc Natl Acad Sci U S A. 2005; 102: 11823-11828.
- Wagnier A, Legros-Maida S, Bosselut R, Bourge JF, Lafaurie C, Ghysdael CJ *et al.* *Identification of human granzyme B promoter regulatory elements interacting with activated T-cell-specific proteins: implication of Ikaros and*

CBF binding sites in promoter activation. Proc Natl Acad Sci U S A. 1995; 92: 6930-6934.

- Wasserman R, Li YS, Hardy RR. *Differential expression of the blk and ret tyrosine kinases during B lineage development is dependent on Ig rearrangement.* J Immunol. 1995; 155: 644-651.
- Watford WT, Hissong BD, Bream JH, Kanno Y, Muul L, O'Shea JJ. *Signaling by IL-12 and IL-23 and the immunoregulatory roles of STAT4.* Immunol Rev. 2004; 202: 139-156.
- Webb R, Wren JD, Jeffries M, Kelly JA, Kaufman KM, Tang Y *et al.* *Variants within MECP2, a key transcription regulator, are associated with increased susceptibility to lupus and differential gene expression in patients with systemic lupus erythematosus.* Arthritis Rheum. 2009; 60: 1076-1084.
- Wellcome Trust Case Control Consortium. *Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls.* Nature. 2007; 447: 661-678.
- Wen ZK, Xu W, Xu L, Cao QH, Wang Y, Chu YW *et al.* *DNA hypomethylation is crucial for apoptotic DNA to induce systemic lupus erythematosus-like autoimmune disease in SLE-non-susceptible mice.* Rheumatology (Oxford). 2007; 46: 1796-1803.
- Wertz IE, O'Rourke KM, Zhou H, Eby M, Aravind L, Seshagiri S *et al.* *De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF-kappaB signalling.* Nature. 2004; 430: 694-699.
- Willcocks LC, Lyons PA, Clatworthy MR, Robinson JI, Yang W, Newland SA *et al.* *Copy number of FCGR3B, which is associated with systemic lupus erythematosus, correlates with protein expression and immune complex uptake.* J Exp Med. 2008; 205: 1573-1582.
- Wilson WA, Gharavi AE, Koike T, Lockshin MD, Branch DW, Piette JC *et al.* *International consensus statement on preliminary classification criteria for*

definite antiphospholipid syndrome: report of an international workshop. Arthritis Rheum. 1999; 42: 1309-1311.

- Wu H, Cantor RM, Graham DS, Lingren CM, Farwell L, Jager PL *et al.* *Association analysis of the R620W polymorphism of protein tyrosine phosphatase PTPN22 in systemic lupus erythematosus families: increased T allele frequency in systemic lupus erythematosus patients with autoimmune thyroid disease.* Arthritis Rheum. 2005; 52: 2396-2402.
- Yamaoka K, Saharinen P, Pesu M, Holt VE 3r, Silvennoinen O, O'Shea JJ. *The Janus kinases (Jaks).* Genome Biol. 2004; 5: 253.
- Yan B, Ye S, Chen G, Kuang M, Shen N, Chen S. *Dysfunctional CD4+ CD25+ regulatory T cells in untreated SLE, secondary to IFN α producing antigen presenting cells.* Arthritis Rheum. 2008; 58: 801-812.
- Yang JQ, Saxena V, Honglin Xu H, Kaer LV, Wang CR, Singh RR. *Repeated α -galactosylceramide administration results in expansion of NK T cells and alleviates inflammatory dermatitis in MRL-lpr/lpr mice.* J Immunol. 2003; 171: 4439-4446.
- Yang W, Ng P, Zhao M, Hirankarn N, Lau CS, Mok CC *et al.* *Population differences in SLE susceptibility genes: STAT4 and BLK, but not PXX, are associated with systemic lupus erythematosus in Hong Kong Chinese.* Genes Immun. 2009; 10: 219-226.
- Yochum GS, McWeeney S, Rajaraman V, Cleland R, Peters S, Goodman RH *et al.* *Serial analysis of chromatin occupancy identifies beta-catenin target genes in colorectal carcinoma cells.* Proc Natl Acad Sci U S A. 2007; 104: 3324-3329.
- Yokoyama K, Su IH, Tezuka T, Yasuda T, Mikoshiba K, Tarakhovsky A *et al.* *BANK regulates BCR-induced calcium mobilization by promoting tyrosine phosphorylation of IP3 receptor.* EMBO J. 2002; 21: 83-92.

- Yung R, Kaplan M, Ray D, Schneider K, Mo RR, Johnson K *et al.* Autoreactive murine Th1 and Th2 cells kill syngeneic macrophages and induce autoantibodies. *Lupus*. 2001; 10: 539-546.
- Yung RL, Quddus J, Chrisp CE, Johnson KJ, Richardson BC. *Mechanism of drug-induced lupus. I. Cloned Th2 cells modified with DNA methylation inhibitors in vitro cause autoimmunity in vivo.* *J Immunol*. 1995; 154: 3025-3035.
- Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T *et al.* *Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes.* *Nat Genet*. 2008; 40: 638-645.
- Zervou MI, Mamoulakis D, Panierakis C, Boumpas DT, Goulielmos GN. *STAT4: A risk factor for type 1 diabetes?* *Hum Immunol*. 2008^a; 69: 647-650.
- Zervou MI, Sidiropoulos P, Petraki E, Vazgiourakis V, Krasoudaki E, Raptopoulou A *et al.* *Association of a TRAF1 and a STAT4 gene polymorphism with increased risk for rheumatoid arthritis in a genetically homogeneous population.* *Hum Immunol*. 2008^b; 69: 567-571.
- Zhao S, Long H, Lu Q. *Epigenetic Perspectives in Systemic Lupus Erythematosus: Pathogenesis, Biomarkers, and Therapeutic Potentials.* *Clin Rev Allergy Immunol*. 2009. [Epub ahead of print].
- Zhernakova A, van Diemen CC, Wijmenga C. *Detecting shared pathogenesis from the shared genetics of immune-related diseases.* *Nat Rev Genet*. 2009; 10: 43-55.
- Zou X, Qiu G, Chen C, Wu M, Hu Y, Zheng H *et al.* *Expression pattern and subcellular localization of five splice isoforms of human PXX.* *Int J Mol Med*. 2005; 16: 701-707.

ANEXO

Lack of replication of genetic predictors for the rheumatoid arthritis response to anti-TNF treatments: a prospective case- only study

Marian Suarez-Gestal^{1*}, Eva Perez-Pampin^{1*}, Manuel Calaza¹, Juan J. Gomez-Reino^{1,2}
and Antonio Gonzalez¹

¹ Laboratorio Investigacion 10 and Rheumatology Unit, Hospital Clinico Universitario
de Santiago, Santiago de Compostela. Spain; mariansg82@hotmail.com,
evappampin@gmail.com, mcalaza@gmail.com

² Department of Medicine. University of Santiago de Compostela, Santiago de
Compostela. Spain; Juan.Jesus.Gomez-Reino.Carnota@sergas.es

* These authors contributed equally

Corresponding author:

Antonio Gonzalez

Laboratorio de Investigacion 10

Hospital Clinico Universitario de Santiago

Travesia de Choupana sn.

15706-Santiago de Compostela

Spain

Tfn: 34 981 950 903

Fax: 34 981 951 068

Antonio.Gonzalez.Martinez.Pedrayo@sergas.es

anlugon@hotmail.com

Keywords: Rheumatoid arthritis, anti-TNF, Gene polymorphism, DAS28, Biomarker

ABSTRACT

Introduction: We aimed to replicate the strong associations that a recent Genome Wide Association Study (GWAS) has found between 16 SNPs and response to anti-TNF treatment in 89 patients with RA. This study is very important because, according with published simulations, associations as strong as the reported will mean that these SNPs could be used as predictors of response at the individual level.

Methods: Disease activity score (DAS28) was evaluated in 151 anti-TNF treated patients with RA of Spanish ancestry at baseline and every 3 months thereafter. Genotypes of the 16 putative predictor SNPs were obtained by single-base extension. Association between the relative change in DAS28 and SNP genotypes was tested by linear regression. In addition, logistic regression was applied to compare genotypes in non-responders (n = 34) versus good-responders (n = 61) following the EULAR response criteria.

Results: None of the analyses showed any significant association between the 16 SNPs and response to anti-TNF treatments at 3 or 6 months. Results were also negative when only patients treated with Infliximab (66.9 % of the total) were separately analyzed. These negative results were obtained in spite of a very good statistical power to replicate the reported strong associations.

Conclusion: We still do not have any sound evidence of genetic variants associated with RA response to anti-TNF treatments. In addition, the possibility we had envisaged of using the results of a recent GWAS for prediction in individual patients should be dismissed.

INTRODUCTION

Anti-tumor necrosis factor (TNF) therapies have revolutionized the treatment of rheumatoid arthritis (RA)[1-2]. Three drugs of this type, Infliximab, Etanercept and Adalimumab, have been used with success in hundreds of thousands of patients with RA around the world. New drugs targeting TNF are in development or have been recently approved[3]. The beneficial effects of these drugs include a better quality of life, control of inflammation, stiffness and pain, and slowing progression to joint erosions and deformity. It seems also that they are able to decrease cardiovascular risk and overall mortality of patients with RA[4-5]. However, there is a significant percentage of patients that do not obtain these advantageous effects[1-3]. In some of them, this lack of response is primary, from the start of the treatment, whereas others develop resistance to treatment after a period of initial response. Unfortunately, there are no useful predictors to forecast how the clinical response of a specific patient will be. This has led to an unsatisfactory trial and error approach in the selection of drugs, which means that some patients will miss an effective treatment at a critical window of opportunity time[6] and a waste of resources for the health services. As a response to this challenge, multiple lines of research are looking for predictors of response to anti-TNF therapies among patient clinical features, synovial tissue biomarkers, blood proteins or genetic variants[7-10]. Very promising, although preliminary, findings have been reported in this last field. Sixteen SNPs with an important association with response to treatment were identified in a recent Genome Wide Association (GWA) study[7]. In our view, the most remarkable aspect of these findings was the marked effect size of each SNP, with levels very rarely found in genetic studies of complex traits. All showed O.R. over 3.5 in the comparison between patients with good response and non-responders. Some of them showed effect sizes over O.R. = 10. These effects together with minor allele

frequencies over 12 % mean that they will allow, if confirmed, for prediction of response to anti-TNF treatments with great accuracy at the level of the individual patient[11]. The limitation of this study was that only 89 patients were included and even very significant results in a study of this size are uncertain. Our objective has been to provide the necessary replication to these exciting findings with the expectation that, at least, a few of them will be confirmed. This will be a first step before proceeding to prospective clinical studies to assess their utility in clinical practice.

MATERIALS AND METHODS

Patients: A group of 151 patients with RA were followed prospectively at the Rheumatology Unit of the “Hospital Clinico Universitario de Santiago” to study the efficacy of anti-TNF therapy. All of them were of European (Spanish) ancestry. Only patients that were naïve with respect to biologic treatments were included. Patients have been systematically evaluated at the initiation of therapy and every three months thereafter. Evaluations included painful and swollen joint counts, visual analogue scales of pain, global health assessment by the patient and the physician, erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), Health Assessment Questionnaire (HAQ) and DAS-28 score. Clinical characteristics are detailed in Table 1. All participants gave their informed consent for inclusion and the study and procedures have been approved by the Ethics Committee for Clinical Research of Galicia.

Assessment of the efficacy of the treatment: We used the same procedures described in Liu et al. to make our results comparable[7]. Response to anti-TNF treatments was assessed with the DAS28 score.[12] The primary outcome was the quantitative variable $relDAS28$, which was calculated as follows:

$$relDAS\ 28 = \left[\frac{DAS\ 28_{baseline} - DAS\ 28_{3\ months}}{DAS\ 28_{baseline}} \right]$$

A secondary outcome was the EULAR response classification in good-, moderate- or non-responders.[13] Good-responders have a $\Delta\text{DAS28} \geq 1.2$ and DAS28 at 3 months ≤ 3.2 ; moderate-responders have $\Delta\text{DAS28} \geq 1.2$ and DAS28 at 3 months > 3.2 or $0.6 < \Delta\text{DAS28} \leq 1.2$ and DAS28 at 3 months ≤ 5.1 ; and non-responders are those who do not fall in any of these categories.

Genotypes: A total of 16 SNPs from Liu et al. [7] were analyzed (Table 2). Genotypes were obtained by single-base extension with the SNaPshot Multiplex Kit (Applied Biosystems, Foster City, CA, USA) and specific primers and probes (available as supplementary material). Genotype call rate was 99.79 %, allele frequencies were in Hardy-Weinberg equilibrium and concordant results for the 16 SNPs were obtained in the 21 samples that were genotyped twice.

Statistical Analysis: Comparisons of the clinical characteristics of the patients with RA included in the GWAS and in our study was done with the Student t test for data available as means and standard deviation and with the Chi squared test for contingency tables for frequency data. Analyses of the relationship between SNPs and treatment response were done as in Liu et al. to make our results comparable also in this aspect[7]. Briefly, linear regression analysis between genotype data following a genetic additive model and relDAS28 as the continuous dependent variable was done. A t statistic was derived from the linear regression and used to calculate the p-value of the association. This statistic is robust to deviations from normality of relDAS28 . We also conducted logistic regression analysis between the groups of responders and non-responders. Odds ratios (O.R.) and their 95 % confidence intervals (C.I.) were obtained using as reference the non-responder group. This second analysis will be less powerful because the phenotype is transformed to a dichotomous variable and because the sample size is reduced by exclusion of the moderate responders. Statistical analyses were performed in

a customized version of the Statistica 7.0 program (StatSoft, Tulsa, OK, USA). We also explored visually the possibility that consideration of all the SNPs jointly will discriminate between responder and non-responder patients. This analysis was done with the Co-Plot algorithm implemented in the Visual Co-Plot software (D. Talby and A. Raveh; available at <http://www.cs.huji.ac.il/~davidt/vcoplot/index.html>) [14]. Estimation of the statistical power for the linear regression analysis was done by transforming the reported P values and number of samples in the corresponding correlation coefficients (R^2). The values of R^2 and the number of samples in our study were used as input in the module for the F test in omnibus comparisons by linear regression of the G*Power v3.0.10 software[15].

RESULTS

The aim of our study was to replicate the strong association of 16 SNPs with response to anti-TNF therapy reported in a recent GWAS[7]. Therefore, we have used the same variables and type of analysis. Data from the 151 patients with RA are shown in table 1. Some of the characteristics of our study population were different from the patients analyzed in the GWAS[7]. Specifically, our patients showed a lower percentage of RF positivity, higher positivity for anti-CCP antibodies, and higher baseline HAQ and DAS28 levels. There were 70.2 % patients with high disease activity at baseline as assessed by a DAS28 > 5.1. In spite of this high activity, there were 40.4 % and 43.7 % good responders at 3 and 6 months, respectively, and only 22.5 % and 21.8 % of non-responders at 3 and 6 months, respectively. The percentages of responders and non-responders were similar in the two studies. In contrast, the proportion of patients treated with each of the three anti-TNF drugs was different (Table 1). In our cohort, most patients were treated with Infliximab (66.9 %), followed by Etanercept (23.2 %) and

Adalimumab (9.9 %). We also checked that there was a good correlation between the variable used as primary outcome in our analysis, redDAS28, and the EULAR response classification (Figure 1) allowing for consistency in the analyses.

The relationship between the SNP genotypes and response to anti-TNF treatment at 3 months was evaluated by linear regression analysis between the genotypes and the continuous variable relDAS28, which is the relative change in DAS28 between baseline and the time of evaluation. There was not association of any of the 16 SNPs with relDAS28 at 3 months (Table 2). Secondary analyses showed very similar results. Comparison of non-responders with good-responders according to the EULAR criteria at 3 months did not show any significant association (Table 2). The most extreme O.R. = 1.9 (95 % C.I. = 1.0-3.6) corresponded to rs437943 in the *CNTDE1* locus, but it compared poorly with the previously reported O.R. = 4.6 (95 % C.I. = 1.8-12.3). In addition, there was not association of relDAS28 with any of the 16 SNPs at 6 months, neither of the classification in responders and non-responders (available as supplementary table). Finally, analysis of response to treatment with Infliximab, which represented 66.9 % of the patients in our study, did not show any significant association (available as supplementary table). No separate analyses of response to treatment with Etanercept or Adalimumab were done because of the small number of patients in these two subgroups. We also visually explored if joint consideration of the 16 SNPs was able to discriminate between the different groups of patients according to their response to treatment, but patients with different responses did not show any clustering in identifiable groups in this analysis (Figure 2).

It was critical, for interpretation of the above results, to assess if our study had enough statistical power to replicate the previously reported associations. Power for the weakest association in the GWAS, which corresponds to rs928655 ($P = 3 \times 10^{-5}$), was

larger than 95 % for a P value of 0.002. It is important to remark that O.R. from the GWAS are very likely heavily biased upwards as consequence of the winner's curse affecting any GWAS and specially those of small size [16-17]. Therefore this power estimate is only valuable in the context of the reported O.R. taken at face value.

DISCUSSION

There is a great need of good predictors for RA response to the anti-TNF treatments[1-3, 9]. The development and approval of new effective drugs for RA adds to this urgency[3]. The recent GWAS from Liu et al. [7] was especially remarkable because it showed such strong associations that, according to published simulations[11], they could be used for prediction in individual patients. This is a characteristic that has not been found in any of the previous studies. However, the size of the study implied that results should be replicated before they could be taken at face value, as already acknowledged by the authors. We have tried to provide here the needed replication in the expectation that some of them will be confirmed and that validation in prospective studies will soon follow.

Unfortunately, none of the associations was replicated in spite of the moderately larger sample size of our study and the corresponding very good power to detect this type of strong associations. These results make very unlikely that any of the 16 SNPs could have an association as strong as suggested by the previous GWAS[7]. It is possible that the differences between the patients with RA in the two studies could have had an effect in the lack of replication, but these differences were not large enough to completely explain the very divergent results. In addition, patients in the GWAS were predominantly of European ancestry as were all the patients in our study. Therefore, it seems more likely that the original strong associations were due to random variation of

allele frequencies in a study including more than 300 000 SNPs and to the heavy bias characteristic of GWAS of small sample size [16-17]. This possibility was already considered by us before beginning this study, but we judged that some SNPs will be replicated given that they showed low P values, 5 of them with P values below 10^{-6} [7], and low P values are the best indication of the reproducibility of results[18].

CONCLUSIONS

Our negative results implicate that we do not have still any strong evidence supporting a significant role of genetic variation in the response to anti-TNF. In addition, our results imply that none of the SNPs in our study will be useful as individual predictors of response to anti-TNF, as we initially expected. It is still possible that they will show association with response to anti-TNF, but of a much weaker strength.

ABBREVIATIONS

GWAS: Genome wide association study

RA: Rheumatoid arthritis

DAS28: Disease activity score

Anti-TNF: Anti-tumor necrosis factor

ESR: Erythrocyte sedimentation rate

CRP: C-reactive protein

HAQ: Health assessment questionnaire

O.R.: Odds ratio

C.I.: Confidence interval

MAF: Minor allele frequency

COMPETING INTERESTS

Roche Spain contributed to the funding of this project. However, the company had no input in the design of the study or in the analysis and writing of the manuscript. The company did not have the right to early access to results or to interfere in any other way with the interpretation or reporting of the results. Therefore, the authors keep exclusive and complete responsibility for the study.

AUTHORS' CONTRIBUTION

MS-G participated in design of the study, genotyped the samples, participated in the interpretation of the results and in writing the manuscript. EP-P participated in the acquisition of clinical data and collection of samples and in the analysis and interpretation of results. MC participated in the statistical analysis and in the interpretation of results. JJG-R coordinated the acquisition of clinical data and participated in the analysis and interpretation of results. AG participated in the design of the study and in the coordination of acquisition of clinical data and collection of samples, supervised genotyping, statistical analysis, interpretation of results and writing of the manuscript. All authors read and approved the final manuscript

ACKNOWLEDGMENTS

We thank Carmen Pena-Pena for her excellent technical assistance and Yolanda Lopez-Golan for her help in recruiting patients. M.S-G is the recipient of a FPU pre-doctoral bursary of the Spanish Ministry of Education; M.C. is the recipient of an "Isabel Barreto" bursary of the Government of Galicia. This project was supported by an unrestricted grant from Roche Spain and by grants PI080744 and PI09/90744 from

the Instituto de Salud Carlos III (Spain) with participation of funds from FEDER (European Union).

REFERENCES

1. Chen YF, Jobanputra P, Barton P, Jowett S, Bryan S, Clark W, Fry-Smith A, Burls A: **A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness.** *Health Technol Assess* 2006, **10**(42):iii-iv, xi-xiii, 1-229.
2. Moreland LW, Emery P: **TNF[alpha]-inhibition in the treatment of rheumatoid arthritis.** London: Martin Dunitz; 2003.
3. Rubbert-Roth A, Finckh A: **Treatment options in patients with rheumatoid arthritis failing initial TNF inhibitor therapy: a critical review.** *Arthritis Res Ther* 2009, **11 Suppl 1**:S1.
4. Carmona L, Descalzo MA, Perez-Pampin E, Ruiz-Montesinos D, Erra A, Cobo T, Gomez-Reino JJ: **All-cause and cause-specific mortality in rheumatoid arthritis are not greater than expected when treated with tumour necrosis factor antagonists.** *Ann Rheum Dis* 2007, **66**(7):880-885.
5. Dixon WG, Watson KD, Lunt M, Hyrich KL, Silman AJ, Symmons DP: **Reduction in the incidence of myocardial infarction in patients with rheumatoid arthritis who respond to anti-tumor necrosis factor alpha therapy: results from the British Society for Rheumatology Biologics Register.** *Arthritis Rheum* 2007, **56**(9):2905-2912.
6. Finckh A, Liang MH, van Herckenrode CM, de Pablo P: **Long-term impact of early treatment on radiographic progression in rheumatoid arthritis: A meta-analysis.** *Arthritis Rheum* 2006, **55**(6):864-872.

7. Liu C, Batliwalla F, Li W, Lee A, Roubenoff R, Beckman E, Khalili H, Damle A, Kern M, Furie R, Dupuis J, Plenge RM, Coenen MJ, Behrens TW, Carulli JP, Gregersen PK: **Genome-wide association scan identifies candidate polymorphisms associated with differential response to anti-TNF treatment in rheumatoid arthritis.** *Mol Med* 2008, **14**(9-10):575-581.
8. Maxwell JR, Potter C, Hyrich KL, Barton A, Worthington J, Isaacs JD, Morgan AW, Wilson AG: **Association of the tumour necrosis factor-308 variant with differential response to anti-TNF agents in the treatment of rheumatoid arthritis.** *Hum Mol Genet* 2008, **17**(22):3532-3538.
9. Skapenko A, Prots I, Schulze-Koops H: **Prognostic factors in rheumatoid arthritis in the era of biologic agents.** *Nat Rev Rheumatol* 2009, **5**(9):491-496.
10. Plenge RM, Criswell LA: **Genetic variants that predict response to anti-tumor necrosis factor therapy in rheumatoid arthritis: current challenges and future directions.** *Curr Opin Rheumatol* 2008, **20**(2):145-152.
11. Janssens AC, Moonesinghe R, Yang Q, Steyerberg EW, van Duijn CM, Khoury MJ: **The impact of genotype frequencies on the clinical validity of genomic profiling for predicting common chronic diseases.** *Genet Med* 2007, **9**(8):528-535.
12. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL: **Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis.** *Arthritis Rheum* 1995, **38**(1):44-48.
13. van Gestel AM, Prevoo ML, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL: **Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with**

- the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. *Arthritis Rheum* 1996, **39**(1):34-40.
14. Bravata DM, Shojania KG, Olkin I, Raveh A: **CoPlot: a tool for visualizing multivariate data in medicine.** *Stat Med* 2008, **27**(12):2234-2247.
 15. Faul F, Erdfelder E, Lang AG, Buchner A: **G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences.** *Behav Res Methods* 2007, **39**(2):175-191.
 16. Ioannidis JP: **Why most discovered true associations are inflated.** *Epidemiology* 2008, **19**(5):640-648.
 17. Xiao R, Boehnke M: **Quantifying and correcting for the winner's curse in genetic association studies.** *Genet Epidemiol* 2009, **33**(5):453-462.
 18. Todd JA: **Statistical false positive or true disease pathway?** *Nat Genet* 2006, **38**(7):731-733.

FIGURE LEGENDS

Figure 1: Good correlation between outcome variables, EULAR response classification and reIDAS28. Median, interquartile range, and non-outlier range are represented as dots, box and whiskers, respectively. Empty dots represent outliers.

Figure 2: Multivariate visual analysis showing that 16 SNPs were not able to separate RA patients classified according to their EULAR response. Responders are represented as blue dots; moderate responders as red dots; and non-responders as green dots. Yellow arrows represent the genotypes of each of the 16 SNPs according to an additive model. This representation was obtained with Visual Co-Plot.;

Table 1. Clinical characteristics of the patients in this study compared with the report from Liu et al. [7]

Variable	This study	Data from ref. 7	P-value
Age (mean \pm SD)	54.2 \pm 13.1	57 \pm 13.5	0.06
Age at diagnosis (mean \pm SD)	46.2 \pm 13.5	47 \pm 15	0.3
Women (%)	84.8	75	0.06
Disease duration (mean \pm SD)	7.8 \pm 6.6	8 \pm 8	0.4
Current Smokers (%) ^a	9.1	15	0.3
Ever Smokers (%) ^a	25.0		
RF+ (%)	68.9	83.8	0.01
CCP+ (%) ^a	86.3	61.9	0.0001
SE+ (%) ^a	58.3		
ANAs+ (%) ^a	29.8		
HAQ at baseline (mean \pm SD)	1.4 \pm 0.7	1.1 \pm 0.6	0.0005
DAS28 at baseline (mean \pm SD)	5.6 \pm 1.2	5.2 \pm 0.8	0.002
DAS28 \geq 5.1 (%)	70.2		
3.2 \leq DAS28 < 5.1 (%)	26.5		
DAS28 at 12-16 wk (mean \pm SD)	3.6 \pm 1.4	3.7 \pm 1.3	0.2
DAS28 at 6 months (mean \pm SD)	3.5 \pm 1.4		
Good-responders at 12-16 wk (%)	40.4	34.8	0.6
Non-responders at 12-16 wk (%)	22.5	25.8	0.6
Good-responders at 6 months (%)	43.7		
Non-responders at 6 months (%)	21.8		
anti-TNF drug			
Infliximab (n)	101	32	0.01
Etanercept (n)	35	39	0.02
Adalimumab (n)	15	18	0.05

^a Data was available for 102 patients for CCP antibodies, ANA and SE; and for 88 patients for smoking habits.

Table 2. Relationship of relDAS28 and SNP genotypes, and comparison of the allele frequencies between responders and non-responders.

SNP	P value of relDAS28	Responders' MAF % (n/N) ^b	Non-responders' MAF % (n/N) ^b	O.R. ^c (95% C.I.)	P value
rs983332	0.6	25.9 (30/116)	27.9 (19/68)	1,11 (0,6-2,0)	0.8
rs928655	0.1	18.9 (23/122)	28.0 (19/68)	0.60 (0.3-1.2)	0.1
rs13393173	0.8	23.8 (29/122)	20.6 (14/68)	0,83 (0,4-1,7)	0.6
rs437943	0.05	38.5 (47/122)	25.0 (17/68)	0,53 (0,3-1,0)	0.06
rs10945919	0.7	27.9 (34/122)	23.5 (16/68)	0,79 (0,4-1,7)	0.5
rs854547	0.3	39.3 (48/122)	36.8 (25/68)	1.12 (0.6-2.1)	0.7
rs854548	0.9	23.0 (28/122)	23.5 (16/68)	1,03 (0,5-2,0)	0.9
rs854555	0.6	35.2 (43/122)	38.2 (26/68)	1,14 (0,6-2,0)	0.7
rs868856	0.6	32.8 (40/122)	36.8 (25/68)	1,19 (0,6-2,0)	0.6
rs7046653	0.5	32.0 (39/122)	36.8 (25/68)	1,23 (0,7-2,5)	0.5
rs2814707	0.8	25.4 (31/122)	29.4 (20/68)	1,22 (0,6-2,5)	0.6
rs3849942	0.4	23.0 (28/122)	29.4 (20/68)	1,41 (0,7-2,5)	0.3
rs774359	0.2	26.2 (32/122)	36.8 (25/68)	1,64 (0,8-3,3)	0.1
rs6138150	0.5	14.0 (17/122)	14.7 (10/68)	0.94 (0.4-2.2)	0.9
rs6028945	0.9	12.3 (15/122)	13.2 (9/68)	1,09 (0,5-2,5)	0.9
rs6071980	0.9	18.9 (23/122)	16.2 (11/68)	0,83 (0,4-2,0)	0.6

^a Major/Minor alleles^b MAF = Minor Allele Frequency, n = number of minor alleles, N = total number of alleles^c O.R. were calculated as in Liu et al [7], taking the allele associated with the non-responders as the numerator of the non-responders odds, and the non-responders odds as the numerator of the O.R.

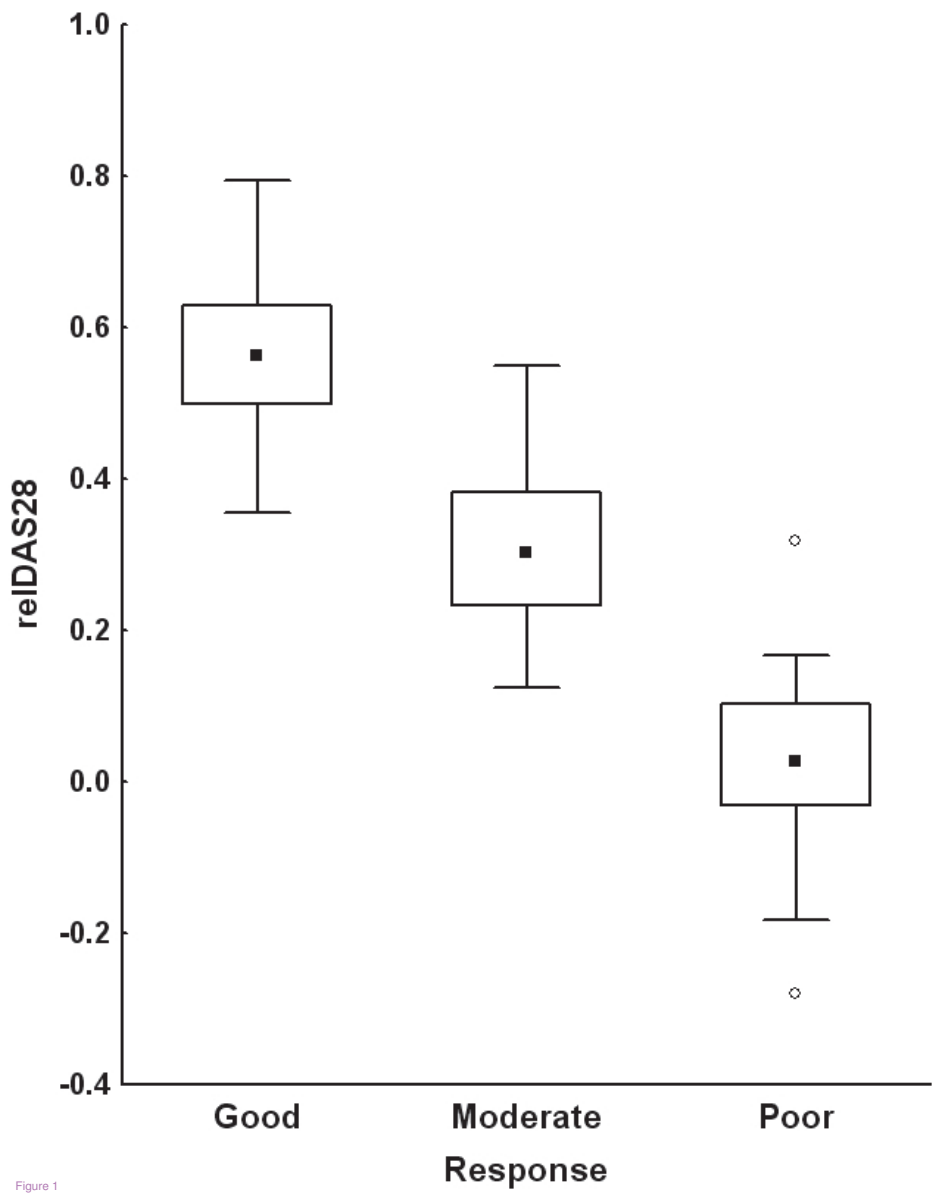
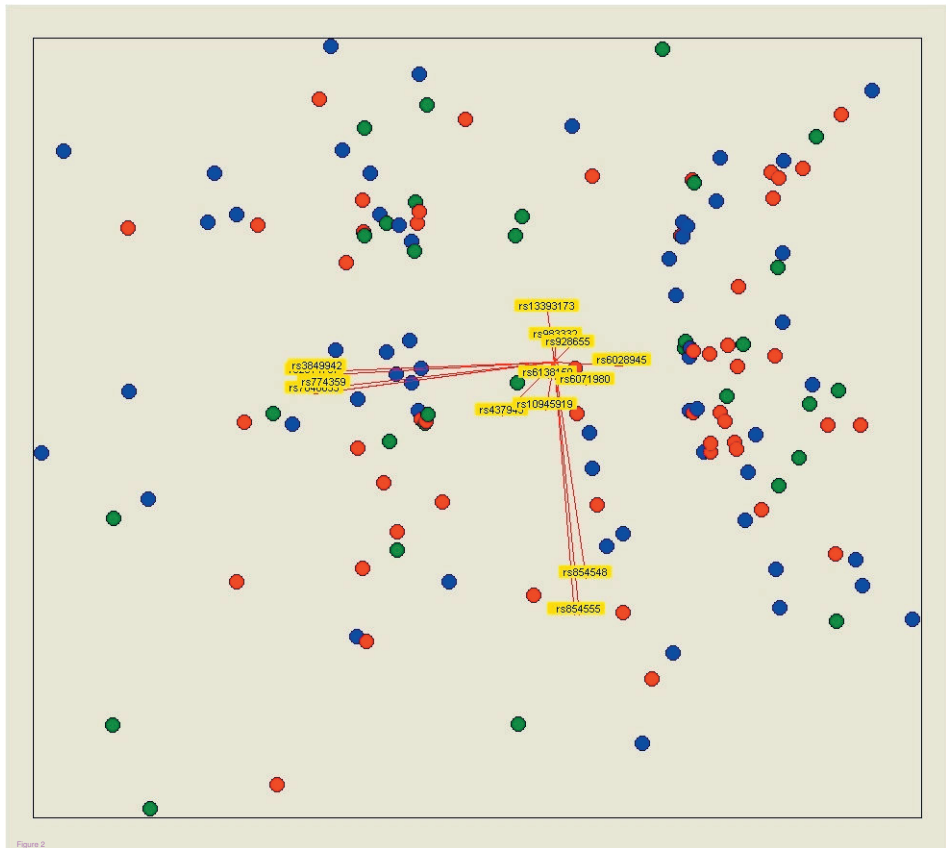


Figure 1

**Additional files provided with this submission:**

Additional file 1: predicting response tto_v07_ART_highlighted changes.doc,
118K

<http://arthritis-research.com/imedia/2026296369357146/supp1.doc>

Supplementary Table 1. Primers and probes in 5' → 3' direction used for genotyping the 16 SNPs by the single-base extension approach.

PCR primers			Minisequencing Probe
SNP	Forward Primer	Reverse Primer	
r#98332	atctgtagctctgtttctgtctgg	caagtactccttctgctgcaaaaag	cagactctctataaattttaaaggaaatcaatacacaaaggtccagcgtccaaacta
r#928655	gcctttcccatcacattctc	ttctcagctgcctcctctctt	CCAGAAGTATATTAATGAGCAGTGGTgaattctgtgatttctcacctgcttccagatttg \$
r#13393173	tgggtgtctcattgctcttgaatg	gaacaagatgggtctctatccccagg	tggaaaagccctCagacagctacaatttagctaaactcca
r#437943	ccagcatcgggtggaagtgat	tggtcagtttgaaaaggtctgc	TTACTATGATTGATgagcaccatatagaacttatagcaacca \$
r#10945919	tcaactgcagttctgcctttg	tgatccactttgtgtttgga	tttggaaaatttactattgaaagagaaaaggttcaaaagctgaaacatttagctgcccaggactcca
r#854547	aagtggtcctcctgcttcag	ttggaccaagactagatttctctg	gctgcttttcaggagtgatgaatcactgaag
r#854548	aacaagccttgactgggtaaac	tgacgggtttgctcttttaac	TTACTATGATgactgggttaaactgfatcaattgtcaaaagGataatctgtccaatca \$
r#854555	cccaagagatacagggttttaac	tttgtccacgggtctttattcag	catcataatataatgataatcatcttgacataatttgtctactgctg
r#868856	cgtgtgtctctcatactcac	agcctttgcagagacacactg	TTACTATGATcctagtcogtcttctcttcttcaacttgggttttcttcccttgctatttaagcatcttccaact \$
r#7046653	tgtaggcacaattggtagctt	gagacactaggaccgtttgttg	tccctggatcctcttatttctgaaattaaaaaaagggctgtatga
r#2814707	aaacgacaaagaagacacagc	gcccactcacaactaaagca	aattcctgcatccagtagtgcccaaaagccagtg
r#33849942	ccatgtaggaactgagacac	cccactacagcccatcttc	ctttctcccacaggtctagtagtacctgatttctttt
r#743459	agtaaggaaactaacatgtaggcactc	gtaccctggggcaaggtg	caactatttctctcatctgtaaaatggcaataatagtaaatagtaacctaatgtgt
r#6138150	caggatggaaagctctgtcgaag	agccaagactcaggggtgtgt	gttggcaaatccagattccaaaagctctc
r#6028945	agctggaaactgggtttgttg	gactcagccttctggttgagg	tccaatgacagtagtgagtaaatgagcatgCactaaataaagaacaca
r#6071980	ctgatctcttttctcgtgttgc	tgcagaattccaaggacacaga	agcctgacaggatttttatttATgtttgggggaacagaataactaacagggtcatgttta

Primers selected with Primer3 and Oligos softwares. They were checked to avoid formation of dimers between the primers included in the reaction.

\$ These oligonucleotides were extended with a 5' tail that has not homology with human sequences (these tails are in capital letters)

Supplementary Table 2. Analysis of the association between genotypes at the 16 SNPs and response to treatment with Infliximab after 3 months of follow-up by linear regression with the relative change in DAS28 (redDAS28) and by comparison of responders and non-responders according to the EULAR criteria.

SNP	<i>P</i> value of redDAS28	Responders' MAF % (n/N) ^a	Non-responders' MAF % (n/N) ^a	O.R. (95% C.I.)	<i>P</i> value
rs983332	0.14	21.2 (14/66)	27.3 (12/44)	0.72 (0.3-1.7)	0.46
rs928655	0.85	20.6 (14/68)	18.2 (8/44)	1.17 (0.4-3.1)	0.75
rs13393173	0.86	22.1 (15/68)	25.0 (11/44)	0.85 (0.3-2.1)	0.72
rs437943	0.67	36.8 (25/68)	25.0 (11/44)	1.74 (0.8-4.0)	0.19
rs10945919	0.47	27.9 (19/68)	22.7 (10/44)	1.32 (0.5-3.2)	0.54
rs854547	0.70	35.3 (24/68)	36.4 (16/44)	0.95 (0.45-2.1)	0.91
rs854548	0.81	22.1 (15/68)	25.0 (11/44)	0.85 (0.3-2.1)	0.72
rs854555	0.92	30.9 (21/68)	36.4 (16/44)	0.78 (0.4-1.7)	0.55
rs868856	0.92	41.2 (28/68)	40.9 (18/44)	1.01 (0.5-2.2)	0.98
rs7046653	0.71	39.7 (27/68)	40.9 (18/44)	0.95 (0.4-2.1)	0.9
rs2814707	0.83	30.9 (21/68)	31.8 (14/44)	0.96 (0.4-2.2)	0.92
rs3849942	0.48	27.9 (19/68)	31.8 (14/44)	0.83 (0.4-1.9)	0.66
rs774359	0.17	30.9 (21/68)	43.2 (19/44)	0.59 (0.3-1.3)	0.18
rs6138150	0.78	13.2 (9/68)	18.2 (8/44)	0.69 (0.2-1.9)	0.48
rs6028945	0.39	13.2 (9/68)	9.1 (4/44)	1.53 (0.4-5.3)	0.5
rs6071980	0.48	22.1 (15/68)	13.6 (6/44)	1.79 (0.6-5.0)	0.26

^a MAF = Minor Allele Frequency, n = number of minor alleles, N = total number of alleles

Supplementary Table 3. Analysis of the of the association between genotypes at the 16 SNPs and response to treatment with anti-TNF after 6 months of follow-up by linear regression with the relative change in DAS28 (redDAS28) and by comparison of responders and non-responders according to the EULAR criteria.

SNP	<i>P</i> value of relDAS28	Responders' MAF % (n/N) ^a	Non-responders' MAF % (n/N) ^a	O.R. (95% C.I.)	<i>P</i> value
rs983332	0.30	19.2 (23/120)	26.7 (16/60)	0.62 (0.3-1.4)	0.23
rs928655	0.23	19.4 (24/124)	25.8 (16/62)	0.68 (0.3-1.4)	0.31
rs13393173	0.45	17.7 (22/124)	27.4 (17/62)	0.61 (0.3-1.2)	0.15
rs437943	0.28	40.3 (50/124)	30.6 (19/62)	1.50 (0.8-2.9)	0.21
rs10945919	0.65	28.2 (35/124)	29.0 (18/62)	0.96 (0.5-1.9)	0.91
rs854547	0.89	35.5 (44/124)	35.5 (22/62)	1.00 (0.5-1.9)	1.00
rs854548	0.85	20.2 (25/124)	21.0 (13/62)	0.96 (0.5-1.9)	0.91
rs854555	0.61	33.9 (42/124)	37.1 (23/62)	0.87 (0.5-1.7)	0.66
rs868856	0.56	31.5 (39/124)	33.9 (21/62)	0.89 (0.5-1.7)	0.74
rs7046653	0.41	30.6 (38/124)	33.9 (21/62)	0.86 (0.4-1.7)	0.65
rs2814707	0.69	21.8 (27/124)	24.2 (15/62)	0.88 (0.4-1.8)	0.72
rs3849942	0.45	20.2 (25/124)	24.2 (15/62)	0.80 (0.4-1.6)	0.54
rs774359	0.53	24.2 (30/124)	29.0 (18/62)	0.78 (0.4-1.6)	0.48
rs6138150	0.054	16.1 (20/124)	11.3 (7/62)	1.43 (0.6-3.4)	0.42
rs6028945	0.75	12.1 (15/124)	14.5 (9/62)	0.78 (0.3-2.1)	0.62
rs6071980	0.84	18.5 (23/124)	16.1 (10/62)	1.22 (0.5-3.0)	0.66

^a MAF = Minor Allele Frequency, n = number of minor alleles, N = total number of alleles